Beyond the Blue Planet: the Formation and Dynamical Evolution of Habitable Worlds

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Sebastian Elser

von

Gossau SG

Promotionskomitee

Prof. Dr. Ben Moore (Vorsitz)
Prof. Dr. George Lake
Prof. Dr. Thomas Gehrmann
Prof. Dr. Uros Seljak
Prof. Dr. Ulrich Straumann

Zürich, 2013

Sebastian Elser
Institute for Theoretical Physics
University of Zurich
Winterthurerstrasse 190
CH-8057 Zürich
Switzerland
selser@physik.uzh.ch

Contents

Acknowledgments **7**

Abstract **9**

Zusammenfassung **11**

1. Introduction **13**

2. Theoretical background **17**
2.1. Planetary Systems and Habitability 17
2.1.1. The Solar System 17
2.1.2. The Search for Extrasolar Planetary Systems 18
2.1.3. Properties of Planetary Systems 19
2.1.4. Habitable planets 21
2.2. Planet formation 23
2.2.1. The protoplanetary disk 24
2.2.2. From dust to planetesimals 25
2.2.3. From planetesimals to protoplanets 26
2.2.4. From planetary embryos to terrestrial planets 28
2.2.5. Formation of giant planets 30
2.2.6. Dynamical evolution of young planetary systems 31
2.3. Simulating the formation and evolution of planets 32
2.3.1. Numerical methods 32
2.3.2. Present day simulations 33

3. Paper I: Earth-Moon Systems **37**

4. Paper II: Elemental Abundances **47**

5. Paper III: Stability of Orbits **65**

6. Prospects **79**

Bibliography **83**

Acknowledgments

Foremost, I am especially grateful to my supervisor Ben Moore for making this thesis possible. Always, he provided helpful advice as well as new ideas and gave me the opportunity to work on various topics in planet formation.

I want to thank all my collaborators, especially Joachim Stadel, Simon Grimm, Michael Meyer and Ryuji Morishima for their fruitful contribution to this thesis and for sharing their expert knowledge with me. Special thanks go to Philippe Jetzer as well as Prasenjit Saha and the Zoological Museum of Zurich for providing me chances to present my work to a wider audience. I am greatful to Doug Potter and Jonathan Coles, who spend their time helping me when I was confronted with any computational problem.

I really appriciated to work at the Institute for Theoretical Physics at the University of Zurich and I want to thank all institute member for contributing their part to the pleasant and unique working environment during the last years. Especially, I thank my office mates for inspiring discussions about topics in physics and beyond.

At the end, I want to thank my fiancée and my family for being a great source of motivation during my studies as well as for giving me a good reason for traveling back home every day.

Abstract

This thesis is about different aspects of the formation of terrestrial planets and the evolution of planetary systems. Terrestrial planets are formed in various stages, from the condensation of dust particles to the collisions of protoplanets. Hence, this process spans orders of magnitude in size and lasts millions of years. The final stages of the formation process and the long-term dynamical evolution of planetary systems can be studied in detail with modern N-body simulations. Such simulations provide the basis for better understanding the origin, property and occurrence rate of planets that provide suitable conditions for the emergence and evolution of life. These worlds are called habitable planets.

In the first part of this thesis we show that massive moons orbiting terrestrial planets are not rare. The Earth's comparatively massive Moon has played an important role in the development of life on our planet. Thus, it is worth to study how often a Earth-like planet is hosting a massive satellite. In a large set of N-body simulations, in which terrestrial planets are formed, we identify giant impacts that can potentially induce the formation of a large moon. Then, we estimate the long-term evolution of the planet-satellite systems considering subsequent impacts, spin-orbit resonance and tidal evolution. According to our study, 1 in 12 Earth-like planets hosts a Moon-like satellite, a surprisingly high number given the fact that it is not the most optimistic approach.

In the second part, we study the sensitivity of the estimation of the bulk composition of terrestrial planets on different models and initial conditions. The elemental abundances in the Earth provide the initial conditions for life and clues to the history and formation of the Solar System. The bulk composition of planets is a result of different processes that take place during their formation, which mix material from different regions in the disk. We model the composition of condensed dust particles with condensation equilibrium calculations and trace the collisional growth and the radial mixing of the grown bodies with N-body simulations. We find that the elemental abundances in terrestrial planets depend strongly on the model that describes the thermodynamical structure of the protoplanetary disk. In general, the composition of the inner Solar System can be reproduced, except from the abundance of highly volatile elements and from the composition of Mercury.

In the third and final part, we scratch a slightly different topic and turn to the long-term evolution and dynamical stability of planetary systems. Most of the currently known extrasolar planets are massive Jovian-like planets, since these planets are easier to detect. Actually, theory predicts that low-mass Earth-like planets are much more numerous. Starting from systems containing two or three known planets, we add additional low-mass planets and are interested in their stability. Numerous N-body simulations provide the long-term evolution of the systems. The stability of the orbits of the hypothetical planets as well as their interaction with the known planets indicate on which orbits unknown planets could exist. Finally, we predict the existence of habitable low-mass planets in most of the systems we took into account.

According to our results, we expect some of them being detected in the next several years.

Zusammenfassung

Diese Dissertation befasst sich mit der Entstehung terrestrischer Planeten und der Entwicklung von Planetensystemen. Die Geburt eines terrestrischen Planeten beginnt mit der Kondensation winziger Staubteilchen und endet mit gigantischen Kollisionen zwischen Protoplaneten. Dieser Prozess umspannt unzählige Grössenordnungen und dauert Millionen von Jahren. Moderne N-Körper Simulationen ermöglichen es sowohl speziell die letzten Etappen in der Entstehungsgeschichte eines Planeten wie auch die anschliessende Evolution ganzer Planetensysteme im Detail zu verfolgen. Sie legen somit die Grundlage für ein besseres Verständnis des Ursprungs, der Eigenschaften und der Häufigkeit terrestrischer Planeten im Universum, die die Entstehung und Entwicklung von Leben ermöglichen. Man spricht dabei von habitablen Planeten.

Der erste Teil dieser Dissertation zeigt, dass terrestrische Planeten nicht selten von einem massereichen Satelliten ähnlich dem Mond begleitet werden. Der Mond spielte eine entscheidende Rolle in der Entwicklung von Lebensformen auf unserem Planeten. Somit liegt die Frage nahe, wie häufig solche Erde-Mond Systeme anzutreffen sind. In zahlreichen N-Körper Simulationen, welche die Entstehung der terrestrischen Planten unseres Sonnensystems beschreiben, identifizieren wir Kollisionen zwischen Protoplaneten, die möglicherweise in der Formation eines mondähnlichen Satelliten resultierten. Die weiter Entwicklung der Umlaufbahnen in jedem dieser Planeten-Mond Systeme kann durch Gezeiteneffekte, Spin-Orbit Resonanzen oder auch nachfolgende Kollisionen stark beeinflusst werden. Diese Effekte beziehen wir durch einfache Modelle ein. Schliesslich halten wir fest, dass etwa jeder zwölfte erdähnliche Planet von einem mondähnlichen Satelliten begleitet wird. Diese überraschend hohe Zahl deutet darauf hin, dass ein mondhnlicher Satellit keine Seltenheit ist.

Der nächste Teil der Dissertation beschäftigt sich mit der Bestimmung der chemischen Zusammensetzung terrestrischer Planeten, und damit, wie stark diese von den verwendeten Modellen abhängt. Die chemische Beschaffenheit der Erde legte die Grundlagen für das Leben auf unserem Planeten und liefert Hinweise über die Geschichte unseres Sonnensystems. Die chemische Komposition eines Planeten wird durch verschiedene Prozesse seiner Enstehungsgeschichte bestimmt, wobei Material unterschiedlichster Herkuft vermischt wird. Unter der Annahme von chemischem Gleichgewicht berechnen wir die Zusammensetzung der kondensierten Staubteilchen und verfolgen den weiteren Weg dieser in den Bausteinen der Planeten eingebetten Feststoffe mit N-Körper Simulationen. Dabei stellen wir fest, dass die mit dieser Methode gefundene Zusammensetzung der terrestrischen Planeten stark von dem verwendeten Model, welches die Thermodynamik in der protoplanetaren Scheibe beschreibt, abhängt. Mit einem geeigneten Modell können die Unterschiede in der Komposition der inneren Planeten unseres Sonnensystems mit Ausnahme von stark flüchtigen Elementen erstaunlich genau reproduziert werden. Merkur, wohl aufgrund seiner einzigartigen Entstehungsgeschichte, ist davon aber ausgenommen.

Der letzte Teil der Dissertation geht nun einen Schritt weiter und stellt die Bahnentwicklung und Stabilität von Planetensystemen ins Zentrum. Die meisten extrasolaren Planeten, welche bis anhin gefunden wurden, sind wesentlich massereicher, beziehungsweise grösser als die Erde; dies ist darauf begründet, dass massereiche Planeten leichter zu beobachten sind als Erdähnliche. Es wird aber vermutet, dass kleinere Planeten eigentlich viel zahlreicher sind. Deshalb untersuchen wir bekannte Planetensysteme auf die mögliche Existenz unentdeckter kleiner Planeten. Hierfür wird ein relativ kleiner Planet, ausgehend von unterschiedlichen Umlaufbahnen, dem System hinzugefügt. Die Stabilität seines Orbits sowie die Wechselwirkung mit den bekannten Planeten liefern Hinweise auf die wahrscheinlichste Umlaufbahn eines hypothetischen Planeten in diesen extrasolaren Systemen. Die meisten Systeme, welche wir in unserer Untersuchung einbezogen haben, könnten einen solchen habitablen Planeten beherbergen. Dieses Wissen kann als Richtlinie für zukünftige Beobachtungen von Nutzen sein.

1

Introduction

The formation and evolution of planetary systems is a fascinating process. It starts with small particles of gas and dust and ends with giant bodies of Earth-mass, Jupiter-mass or more orbiting a central star for billions of years. Observations reveal not only details of the planets of our Solar System but also the properties of extra solar planetary systems more than thousands of lightyears away. Given by the nature of observations, they show only very short snapshots of the life of a planet or a planetary system. Some of those far away systems are observed in different stages of the formation process and the continuous improvement of observational instruments provides steadily better insights. Beside observations, simulations are an essential tool to verify and improve current theories about planet formation and the evolution of planetary systems. New codes and the steady increase in computational power and availability of computational resources make it possible to use such simulations as a helpful tool. They connect observations, models and theory.

This dissertation presents the research projects I have been part of during the three and a half years of my PhD studies. It mainly consists of my three first-author papers. The goal of this thesis was to study different aspects of planet formation and planetary system with N-body simulations. Terrestrial planets like the Earth were in the focus of my research. I studied how often massive satellites form via giant impacts, how the estimated bulk composition of planets depends on model parameters and initial conditions and I predicted possible orbits of Super-Earth size planets in extra-solar planetary systems. Although my studies cover different topics, habitable planets, planets that provide surface conditions to maintain liquid water, run like a common thread through all articles.

In the first paper, we tried to answer a well-defined and severe question: how common are Earth-Moon planetary systems? The Moon plays a fundamental role in the emergence and evolution of life on Earth. Its tides have possibly promoted to replication of early bio-molecules in coastal regions and thanks to spin-orbit resonances, its presence prevents Earth's spin axis from tumbling and thus, guarantees a stable

climate on Earth. Starting from N-body simulations designed to reproduce the terrestrial planets of our Solar System, we estimated the frequency of Moon-forming giant impacts. The long-term evolution of the planet-satellite system is governed by angular-momentum transfer, spin-orbit resonance and subsequent impacts. We found that 1 in 12 Earth-mass planets hosts a Moon-mass satellite over a main-sequence star lifetime. Uncertainties in the estimations of the satellites mass and its long-term evolution result in a low-end estimate of 1 in 45 and a high-end estimate of 1 in 4. Nevertheless, this surprisingly high frequency shows that Earth-Moon like systems are not rare.

In the second paper, the bulk composition of terrestrial planets was in our focus. It is well known that the planets of the inner Solar Systems differ in composition. This differences result from various stages of planet formation: from the inhomogenious condensation of solids in the protoplanetary disk, the radial mixing during the formation process via planetesimals and planetary embryos and processes like volatile loss during collisions. We used different models of the thermodynamic structure of the protoplanetary disk, equilibrium condensation calculations and the same N-body simulations as in the first paper to trace the composition of solids from dust particles to protoplanets as consistently as possible. Finally, we found that the bulk composition of the Solar System planets could be reproduced partially, with discrepancies in volatile elements. In fact, the final compositions and their diversity depend on the disk model significantly. For example, the abundance of hydrogen, an element that plays a fundamental role in the emergence and evolution of life on Earth, is not present in all formed planets and thus, different sources of water delivery have to be available.

In the third paper, we studied the long-term evolution of planetary system. We used the dynamical stability of such systems to deduce the existence of additional, not yet observed planets: in known planetary systems with two or three planets we added a hypothetical planet of the mass of a Super-Earth. We simulated the dynamical evolution of the system during 10 Myr. To guarantee a suitable resolution when studying the parameter space of semi-major axis and eccentricity, we carried out thousands of simulations. This was facilitated by a powerful new N-body code, which runs completely on graphic cards. Scattering events, collisions, secular evolution of the orbits and the back reaction of the known planets on the additional one highlight stable regions in between the planets, where potential Super-Earths could orbit. In most of the systems we looked at, these stable islands were located at least partially in the habitable zone, the range in distance to the star where habitable planets can exist. Upcoming observations will reveal if our predictions are correct.

The thesis is structured as follows: first, in chapter 2, we give an overview of different aspects of planetary systems and their formation. Properties of the planetary systems like the Solar System as well as the search for extra-solar planets are briefly summarized. In addition, the stages of terrestrial planet formation from small particles of dust to protoplanets are presented. Numerical tools to study in particular the final steps of planet formation are shown, whereas an exemplary N-body simulation illustrates their strength. This chapter also highlights in which way the papers are linked to the big picture of planet formation. We avoid to give mathematical details of models and methods, because they are often described extensively in the articles.

Next, in chapter 3 to 5, the three above mentioned papers are presented, each of them with a short introduction which stresses especially our motivation to write the paper. Finally, chapter 6 gives an outlook and highlights improvements which could provide a better insight in planet formation.

2

Theoretical background

2.1. Planetary Systems and Habitability

Towards the end of the 20th century, the Solar System was the only known planetary system. Ground breaking discoveries in the past decades revealed that the Sun is not the only star harboring planets. The ongoing discoveries of extrasolar planetary systems provide a growing database which is of great importance when studying the formation and evolution of planetary systems. In addition, it shows how the Solar System differs to other systems and where they show similarities and especially that planets like the Earth are not unique. In the following, we want to briefly present some of these topics. First, we review some characteristics of the Solar System. Then, we describe how extra-solar planets are discovered and summarize their properties. Finally, we discuss which minimum conditions a planet must fulfill to be called a habitable planet.

2.1.1. The Solar System

The Solar System is around 4.57 Gyr old [1], formed out of interstellar gas and dust. While we are discussing its formation in detail in section 2.2, here, we want to point out some of its remarkable characteristics.

More than 99.98% of the total mass of the Solar System are found in the Sun (stellar mass = $1M_{\odot}$). The largest planets are the gas giants Jupiter ($M_{\text{Jupiter}} \approx 10^{-3}M_{\odot}$) and Saturn and the ice giants Uranus and Neptun, which build up to outer part of the Solar System. While Jupiter orbits the Sun at roughly 5 AU (mean distance between the Sun and the Earth = 1 AU), Neptun's orbit is located at approximately 30 AU. There are numerous minor objects gravitationally bound to the Sun but located beyond Neptun's orbit: dwarf planets like Pluto, comets in the Kuiper Belt [2] and in the Oort cloud [3]. The inner Solar System is located inside Jupiter's orbit. It consists of the terrestrial planets Mercury, Venus, Earth and Mars. These planets are confined in the zone between 0.3 AU and 1.5 AU. The

largest terrestrial planet is the Earth with $M_\oplus \approx 3.3 \times 10^{-3} M_{\text{Jupiter}}$. The astroid belt, harboring the dwarf planet Ceres, is located between Mars and Jupiter.

All planets orbit the Sun in the same direction as the Sun rotates. This prograde orbital motion takes place very close to the plane given by the Sun's equator. In addition, most of the planets, except from Venus and Uranus, rotate in the same sense as the Sun. The largest part of the total angular momentum of the Solar System is contained in the orbital motion of the planets. These facts point towards a common birth place of the Sun and its planets and a segregation of mass and angular momentum during their formation.

The bulk composition of the Sun, the gas planets, the ice planets and the terrestrial planets are fundamentally different. While the Sun consists of 99.9% hydrogen and helium by mass, the planets are enriched by metals (elements havier than helium). The planets in the outer system have a thick gaseous envelope and a core composed of metals. The bulk abundance of hydrogen and helium in Jupiter and Saturn is around 80%. This fraction is around 20% in the case of the ice giants. Their composition is dominated by rock or ice, located in the inner parts of the planet. The thin atmospheres of the terrestrial planets do not contribute significantly to their mass. Terrestrial planets are composed primarily of refractory rocky material. The differences in the bulk composition of the planets is mainly a result of the different stages of the formation of the Solar System. An overview concerning the bulk composition of the terrestrial planets is given in chapter 4.

Most planets, except from Mercury and Venus, have natural satellites. The satellites in the outer Solar System are numerous. They build complex systems with a large range in mass and orbital resonances between satellites. Satellites are formed during or after the accretion phase of the gas in a circumplanetary disk [4] or are captured [5]. The most remarkable satellite hosted by the terrestrial planets is the Moon. Similar to Pluto's satellite Charon, it formed out of the debris of a giant impact. Isotope records indicate that this impact took place on Earth around 100 Myr after the formation of the Solar System [6]. The large mass of the Moon ($\approx 1/81 M_\oplus$) provides tidal effects on the Earth as well as stabilisation of the Earth's rotation axis. Therefore its presence is of great importance for the emergence, evolution and survival of life on Earth. This topic is discussed in chapter 3.

2.1.2. The Search for Extrasolar Planetary Systems

The first extrasolar planet (or exoplanet) orbiting a sun-like star has been discovered in 1995 by the Swiss astronomers Michel Mayor and Didier Queloz [7]. Since than, more than 1'000 exoplanets have been detected or wait for confirmation and the number is continuously increasing [8], e.g. thanks to ongoing space missions like KEPLER or high precision ground based instruments like HARPS.

The detection and characterization of exoplanets is a challenging task. Several methods have been established during the last years. We will now highlight two of the most prominent methods. The *radial velocity* method is based on the data obtained from Doppler spectroscopy measurements of the star which is harboring the potential planet. Due to the movement of the star and the planet around their common barycenter, the star moves back and forth along the line of sight inducing a Doppler

signal in the emitted light. The radial velocity signal in the light curve contains information about the mass, the orbital period and eccentricity of the planet [9]. In fact, the estimated mass is the minimum mass and there is a degeneracy between this mass and the orbital inclination. The sensitivity of the radial velocity measurements is impressive. Stellar movements of the order $\sim 1\,\mathrm{m/s}$ can be measured in the best case with current instruments like HARPS [10]. Theoretically, this is sensitive enough to detect the Sun's movement due to Jupiter, but a detection of the effect of the Earth at its current semi-major axis is out of reach yet. Nevertheless, Earth-mass planets very close to their host star can be detected today [11, 12]. Hence, the more massive and the more close the planet, the higher is its chance to be detected via Doppler spectrography.

The *transit* method has become very important during the last years. When a planet's orbit is seen nearly edge-on, the planet transits its host star periodically. Since such an orientation does not occur with a high probability, many systems are excepted from being observed by this method. Nevertheless, it is very powerful. The transit curve provides information about the size and the orbit of the planet. Short-period planet transit curves can provide information on additional long-period planets in the system: the transit-timing models [13] reveal information about unknown planets hidden in a known transit curve. A characterization of the planet is possible concerning atmospheric composition and density, since the light received during transit has partially passed transparent parts of the planet (e.g. [14]). In addition, the transit of a satellite, a so called exomoon, might be detectable [15]. In general, transit observation can help to break the mass-degeneracy of radial velocity measurements. Similar to the radial velocity method, the transit method is limited to relatively short periods since the survey duration is often limited to a few years or less.

There are other methods, which have provided only a small fraction of the known planet to date. In return, they open up new regions in the parameter space. *Gravitational micro-lensing* [16, 17] is a very sensitive technique but the signal can be detected only once. Potentially, it can lead to the detection of planets in other galaxies [18] and it is the only method to detect free-floating planets, i.e. planet-mass objects not bound to a star. The *direct imaging* of exoplanets is very limited at the moment. A large separation between the star and its planet is required and hence, this method focuses on massive planets at large radii. In the future, new instruments will provide higher angular resolution and deliver better insights into the fascinating world of extra solar planets.

2.1.3. Properties of Planetary Systems

The numerous exoplanets that have been discovered during the past two decades provide a huge database about properties of extra solar systems. In figure 2.1, the distribution of the known exoplanets is shown. It is clearly seen that the semi-major axis and minimum mass parameter space is not sampled homogeneously. No planets with low masses and high semi-major axes are found yet. In addition, the different detection methods are sensitive to different parameter regions. Transit measurements provide observations of planets with semi-major axis smaller than

1 AU, whereas Kepler measurements provide a clear improvement towards smaller respectively less massive planets. Especially in the mass range of so called Super-Earths (2–10 $M_\oplus$), Kepler found numerous candidates. Radial velocity detections cover a large region with a focus on Jupiter-mass planets. At the moment, direct imaging give only detections of large mass planets at large separation. Under the assumption that the Solar System is not unique in system architecture and planet masses, a comparison with the Solar System planets shows clearly the incompleteness of the known distribution as a result of observational limitations.

There is no reason to believe that there are planets distributed all over the parameter space. Hence, we want to highlight some results of statistical studies on the planet distribution. Although Jupiter mass planets at small semi-major axis seem to be quiet common, only 1% of all stars host a so called hot Jupiter at $a < 0.1$ AU [19]. Moreover, small planets are more frequent than large planets [20]. Based on Kepler data, it has been shown that every sixth main squence star has a Earth-size planets with a period below 85 days, much smaller than the giant planet occurrence rate, which is around 5% for periods below 400 days [21]. The high number of close-in Super-Earth planet candidates indicates that such planets are common. In fact, more than 50% of all Sun-like stars might harbor a Super-Earth [22]. This could mean that the Solar System, lacking of Super-Earth size planets, is not the rule but and exeption [23].

Multi-planetary systems are common [24]. Similar to the Solar System, they are often close to coplanarity. This is deduced from transit observations, since they provide the inclination intrinsically.

It is remarkable that many observed planets have high eccentricities. The median eccentricity is around 0.2 [25], which is larger than the eccentricity of any planet in the Solar System (expect from Mercury). This indicates that the circular nature of the orbits of the Solar System planets are not common. Most likely, planet-planet interactions like scattering are responsible for the wide distribution of eccentricity in extrasolar systems [25, 26]. Some examples of known planetary systems with relatively high eccentricities are discussed in chapter 5.

Beside rough parameters like mass, size or orbital period, other more subtle characteristics of planets are subject of ongoing research. Fischer and Valenti (2005) [27] pointed out that the fractions of metals in a star is correlated to the planet frequency: they found a smooth and rapid rise in the fraction of stars with planets, if the star's metallicity is higher than solar metallicity. This correlation holds only for Jovian companions but is not found for Neptune-like or terrestrial planets [28].

The atmospheric compositions of Earth-mass planets, Super-Earths or gas giants are of great interest. Their spectral signature could contain indicators of biological activities [29]. Moreover, the study of the climate on such distant worlds and a comparison with our Earth becomes more and more realistic [30]. In the following, we will discuss the properties of a planet that provides in some ways Earth-like conditions — a habitable world.

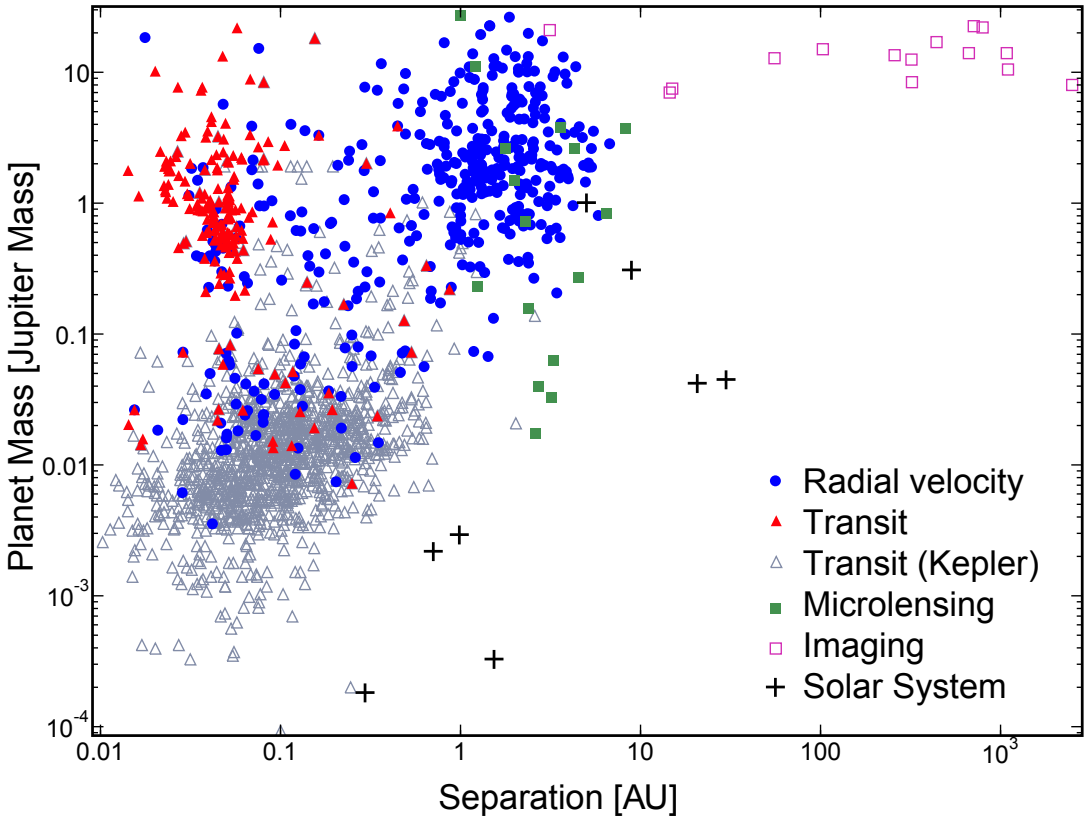

Figure 2.1. The known extrasolar planets. The detection method (the Kepler candidates are shown separately) is indicated by different markers. The semi-major axis vs the mass (or minimum mass) of the planets is shown. The crosses represent the Solar System planets. This plot was generated on 26 February 2013 by exoplanets.org.

2.1.4. Habitable planets

It is the holy grail of exoplanet observations: the detection of a second Earth, orbiting its Sun-like host star inside the so called *habitable zone* (HZ). The detection of an Earth-twin is still outstanding, but it will be an outstanding observation. To date, some recent observations report of Earth-mass planets [31], but they are not in the HZ and are too close to their star, or Super-Earth candidates located in the HZ [32, 33].

The HZ is commonly defined as the region around a star in which water is liquid on a planet's surface [34]. This is motivated by the fact that all lifeforms we know require liquid water during parts of their life cycle.

The radial extension of the HZ depends on different parameters. First of all, it depends on the stellar type. The luminosity and temperature of the star give the energy flux which decreases with distance from the star. Hence, the HZ of a Sun-like G-star is located more distant to the star than the HZ around e.g. a M-dwarf. This circumstance plays to the detectability of planets in the HZ of small stars, since planets are in general easier to detect if they are closer to the star. During their main-sequence lifetime, a star like the Sun increases in luminosity significantly. This results in a continuous shift of the HZ and only a small part of the annulus is always

part of HZ. This part is called the continuous HZ.

As we have seen, the distance to the star gives a rough estimate of the location of the HZ. The properties of the planet, more exactly its atmosphere, can fundamentally alter the HZ. Without an atmosphere, the surface temperature of a planet is given by the solar flux received by the surface, diminished by the albedo of the planet. In the case of the Earth, assuming energy balance and that the planet receives as much energy as it looses via radiation, the surface temperature would be around -18°C, well below the freezing point of water. Fortunately, the actual average temperature is 33°C higher. This is caused by the (in)famous greenhouse effect. Our atmosphere reflects partially the infrared emission of the surface and shifts the energy balance towards higher temperature. Hence, on the one hand, the maximum distance of a habitable planet to the star and hence the outer edge of the HZ is given by the ability of the atmosphere to maintain a greenhouse effect. On the other hand, close to the star, this effect counteracts habitability. Increasing solar insolation results in evaporation of the surface water. At first, this leads to the formation of a cloud cover that holds back the stellar energy release and the inner edge of the HZ is slightly shifted inwards. If the distance to the star is decreased more and more the enhancement of the greenhouse effect results finally in a temperature runaway and the complete evaporation of all water which quickly escapes to space.

Beside a suitable atmosphere and the location of the planet inside the HZ, other factors can be important. If the orbit of the planet has a high eccentricity, it might partially leave the HZ. Nevertheless, this does not have to affect the planet's habitability, since the annual average flux is the critical quantity [35].

A large or chaotic obliquity, the angle of the spin axis relative to the normal of the orbital plane, could result in extreme climates and variations and damp the evolution of life. Hence, it was proposed that a massive satellite might be important for the emergence and evolution of life. As mentioned above, in the case of the Earth-Moon system, the Moon stabilizes the Earth's obliquity [36] and the frequent tides might have boost the emergence of life [37]. Hence, a massive satellite might be important for the habitability of a planet. This hypothesis was the motivation for the work presented in chapter 3.

The existence of liquid water on a planet is inextricably linked to the delivery of water to the planet during or after its formation. As we discuss in chapter 4, water delivery is not guaranteed during the formation process and can depend on rare events like comet impacts.

It is still a long way from providing the essential groundwork for the emergence of life to the evolved diversity of lifeforms we know on Earth. According to the *Rare Earth* hypothesis, a planet as well suited for the emergence and evolution of life like the Earth is very rare, since it depends on numerous factors and rare events [38]. Upcoming detections and observations of potential habitable planets will help to answer the question about the likelihood of other "'inhabited"' planets in our cosmic neighborhood.

2.2. Planet formation

The formation of the Solar System and the formation of planets and stars in general is a field of research that mainly started in the 18th century. Motivated by the nearly circular and coplanar orbits of the Solar System planets, both Immanuel Kant [39] and Pierre-Simon Laplace [40], among others, proposed that planets form in a flat disk that orbits around the central star. In the following centuries, new theories were born, like the close approach of another star that draw matter out of the Sun out of which the planets are formed [41] or the individual agglomeration of the Sun and the planets out of a nebula and a subsequent capture of the planets in the Suns gravitational field [42]. However, all these great ideas have their own drawbacks. It was the Soviet astronomer Victor Safronov [43] who laid the groundwork for many of the fundamental ideas of the modern widely accepted theory in planet formation.

The present day picture of planet formation picks up the early nebular hypothesis. Every star and its planetary systems form during the collapse of a giant molecular cloud. In fact, the cloud fragments into numerous part, each of them a birthplace of a star and its planetary system. Gravity forces the fragment of the cloud to collapse further. Since the angular momentum is conserved, the nebula spins faster and this rotation results in the formation of a flattened disk. In the center of the disk, were most of the mass collects, the temperature increase due to release of kinetic energy and molecular collisions. A protostar is born. A byproduct of the stellar formation process is the circumstellar disk, which contains only a small fraction of the star's mass but most of the angular momentum of the initial nebula as a result of the accretion phase. Observations reveal that such disk can have a diameter more than 10 times larger than the diameter of the current Solar System (e.g. [44]). The formation of this circumstellar disk marks the beginning of the planetary system. Hence, it is often referred as the protoplanetary disk.

Here, we want to review the current theory of planet formation starting from the protoplanetary disk. In the following sections, we discuss the main mechanisms that are involved in the formation of planets in this disk. This is a multi-stage process that begins with gas and dust and ends with planets. As a guideline, figure 2.2 illustrates the formation process using the example of the six inner planets of the Solar System. The given time scale is just a rough estimate and increasing anticlockwise. First (a cut through the initial protoplanetary disk is shown in $t = 0$), when the disk cools, gas condensates into solids, which settle to the disk midplane. The dust particles agglomerate and grow in size. This stage is dominated by interaction with the gaseous disk and electromagnetic forces. Different mechanisms are proposed that allow a growth up to $\sim$km size ($t = 0.1\,\mathrm{Myr}$). In this mass regime, gravity becomes the dominant force. Those bodies are called planetesimals which are still embedded in the protoplanetary gas disk. Planetesimals accrete mass via pairwise mergers with other planetesimals. Larger bodies grow faster than smaller bodies and soon, some grow to the size of planetary embryos which are able to clean their region of dynamical influence from other bodies. This growth depends also on the amount of material that a body can accrete. There exist two regimes: the regime beyond the so called *snow line* allows the condensation of ices and the growth time scale of embryos is very short. After 1 Myr, giant planet cores can form and accrete

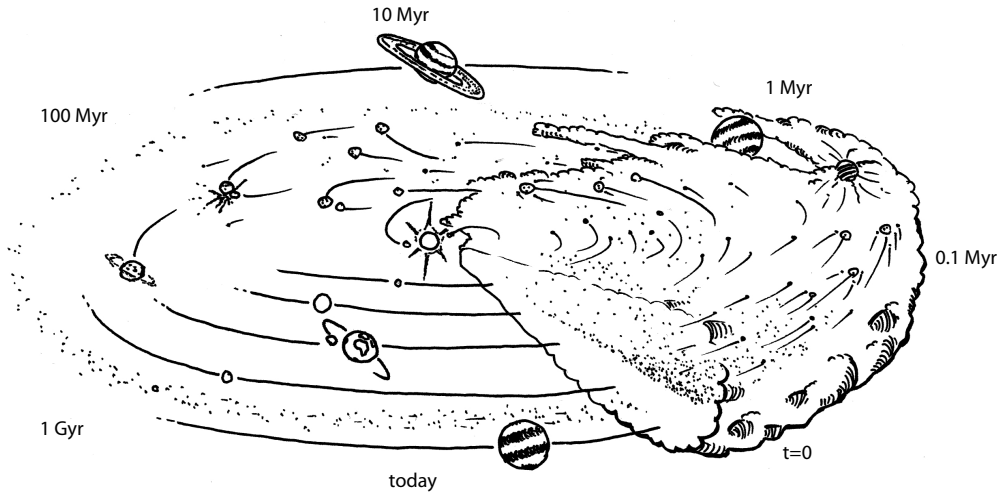

Figure 2.2. A sketch of the formation process of the Solar System up to Saturn. Details are given in the text. (credit: S. Elser)

gas from the disk to become giant planets. In the meantime, inside the snow line, embryos grow slower, while the gas disk dissipates continuously due to radiation pressure and evaporation by the high energetic radiation of the protostar. Secular perturbation by the gas giant, e.g Saturn, result in crossing orbits and the collisions with protoplanets of the similar size (t =1–100 Myr). These giant impacts play a fundamental role in the formation of terrestrial planets like the Earth. A prominent example is the Moon forming impact ($t = 100$ Myr). Finally, we get the present day Solar System.

There are detailed reviews focusing on different topics of terrestrial planet formation, which are the basis for the next chapter. We highlight Lunine et al. (2009) [45] and Morbidelli et al. (2012) [46]. Although 20 years old, Lissauer (1993) [47] still provides one of the best reviews. Last but not least, Armitage (2010) [48] provides an useful overview over numerous topics in planet formation.

2.2.1. The protoplanetary disk

The protoplanetary disk is composed of gas and condensed matter. Its composition is very close to the composition of the central star due to their common formation history. The radiation of the star and energy release of infalling material heats the inner region of the disk and parts of its surface so that interstellar grains are vaporized. In other regions, when the gas infall slows down, the gas can cool and solids start to condense.

The formation of condensates in the protoplanetary disk is a complex process. It is useful to start by assuming that condensates and gas are in equilibrium. This means that the rate of chemical reactions are fast compared to the rate of change in pressure and temperature. In addition, gas and condensates are always in direct contact to provide all possible reactions and the accretion rate of dust particles is slow compared to the reaction rate [49]. Hence, such equilibrium condensation calcula-

tions are path-independent, which makes the calculation straight forward. Given all elements and all species involved in the system, the Gibbs free energy is iteratively minimized at a given temperature and pressure. Thus, it does not depend on the chemical configuration of the system at a earlier state, that means at higher temperature and pressure. In fact, the assumption of equilibrium might provide a good framework to study the composition of solids in the protoplanetary disk. Nevertheless, path-dependent processes likely occur in the disk, resulting from fast accretion or cooling. Since reaction rates are slower and the amount of condensates is high at low temperature, non-equilibrium effects play a significant role at temperature and pressure regimes where volatile elements like hydrogen become part of solids. We discuss the chemistry of a protoplanetary disk in more detail in chapter 4.

Since the condensation sequence depends on the gas composition and pressure, on the temperature and on the change in temperature, the composition of solids in the disk depends very much on the location. Beyond a certain distance to the star (around 2–4 AU), at the snow line, the amount of solid material is much higher than closer to the star [50]. This provides an excellent ground for fast planet formation and promotes the formation of giant planets, as we will discuss later.

A key aspect of planet formation theory is the mass and angular momentum transport inside of the protoplanetary disk. Approximately, ignoring effects like gas pressure support, the angular momentum l per mass unit of the gas in a disk depends on the stellar mass $M_\star$ and on the distance to the star d as follows:

$$l = \sqrt{GM_\star d}, \tag{2.1}$$

where G is the gravitational constant. Thus, l is an increasing function of d. To enable transportation of mass towards the central star, the angular momentum has to be transported from fast rotating inner parts of the disk to slower rotating outer parts. Thanks to this transportation, accretion via a disk is possible. This process and the mathematical description of the disks geometry and evolution is described in detail in chapter 4, or in [48, 51].

2.2.2. From dust to planetesimals

The growth of dust grains proceeds by mutual collisions, whereas the mechanical and chemical processes are not yet completely understood. Dust particles are coupled to the gas, that means they are basically moving with the gas. However, gravity of the star and the missing pressure support forces them to move slightly different depending on their mass. Due to this relative movement of the particles to eachother, particles with different mass collide with velocities with might cause significant fragmentation. In addition, the relative movement of the particles with respect to the gas results in vertical and radial drift.

The vertical component of the star's gravity forces dust particles to settle towards the midplane. During this sedimentation, larger grains move through a background of smaller grains, which are more coupled to the gas and their settling can be neglected. The large particles grow in mass via coagulation until they reach the midplane [43, 52]. Of course, there are other effects like turbulences, Brownian motions and

fragmentation which affect the settlement, but finally, it leads to a enhancement of the large dust grain density at the disk midplane [53].

The radial component of the star's gravity results in a slow radial drift of the dust particles. The orbital movement of the gas is supported by the gas pressure, that means that its orbital velocity is smaller than the Keplerian velocity. When particles grow, they are not perfectly coupled to the gas anymore. Their orbital velocity increases towards the Keplerian velocity and they feel a headwind due to the gas at lower orbital speed. Hence, the particles are slowed down and spiral inwards [54]. Basically, the radial drift is small for very small particles, as they are coupled to the gas, and for very large particles, which have large inertia. Particles with an intermediate size suffer the strongest radial drift. Its magnitude depends on the properties of the gas disk and on the distance to the star. In a typical disk, the radial drift at 1 AU reaches a maximum for meter size particles, which means a drift time scale of $\sim 100\,\mathrm{yr}$. This is much faster than the typical growth rate and results in a catastrophic loss of mass into the star. This is often referred as the *meter-size barrier*.

To overcome the meter-size barrier, the formation of planetesimals through the meter-size regime has to occur very quickly. Ordered pairwise collisions in the regime of cm to km-size bodies is barely understood. In fact, cm-size bodies do not interact significantly by gravity neither stick together by electrostatic forces. Moreover, the large relative drift of particles in this size regime results in high-velocity collision with significant fragmentation [52, 55]. A planetesimal formation with orderly growth might only be possible under certain lucky circumstances. A small fraction of planetesimals might grow fast enough with low-energetic collisions and pass the barrier [56] so that it can grow further. An other possibility to overcome the barrier is a modification of the radial drift that can provide local dust particle enhancement. In general, the radial drift moves particles towards pressure maxima. Typically, the center of the protoplanetary disk is a global maxima, since the pressure decreases outwards. However, local pressure maxima [57] induced by gravitational instabilities [58] or vortices [59] can provide such pressure bumps, where large particles are collected and can grow slowly with low-energy collisions.

An elegant mechanism which can completely ignore the meter-size regime is the Goldreich-Ward mechanism [60]. Vertical settling results in thin disk of high dust density at the disk midplane. When the disk becomes thin enough to be gravitational unstable (classically described by Toomre [61]), planetesimals may form from the fragmentation of the disk in a time scale of order $10^3\,\mathrm{yr}$ [60]. In fact, shear between the high density midplane sub-disk and the low density part of the disk induces Kelvin-Helmholtz instability that make it extremely difficult to get a sub-disk thin enough to reach the instability limit [62, 63]. However, in the midplane, steaming instabilities [64, 65] can provide clumping of cm-sized dust particle. This model can provide a one-stage formation of 1000 km planetary embryos.

2.2.3. From planetesimals to protoplanets

The next stage in planet formation starts when a substantial population of planetesimals has formed. Gravity is the dominant force in this stage and further growth

of the planetesimals occurs via pairwise collisions. These collisions may lead to accretion and fragmentation. The interaction of the solids with the gas disk is not as important as before. Nevertheless, gas drag plays a crucial role in damping inclinations and eccentricities of the bodies and the gas disk provides migration in the early evolution of planetary systems. This is shown in section 2.2.6.

The local initial distribution of planetesimals depends on the distance to the star and on the properties of the protoplanetary disk. Depending on the formation mechanism, planetesimals of 10–1000 km in size may occur, whereas the largest planetesimals might be located beyond the snow line. The planetesimals on Keplerian orbits are perturbed by physical collisions and gravitational scattering and some of the Keplerian motion turns into random motion. The velocity dispersion becomes greater for smaller bodies than for larger bodies as a result of dynamical friction [66, 67]. The further evolution of the planetesimal distribution is dominated by the rapid growth of the largest particles. This is called *runaway growth* and can be explained as follows.

In general, the collisional cross section of a body is larger than its geometrical cross section. Gravity changes the trajectories of bodies that come close and enhances the collisional cross section depending on the masses m_1 and m_2, sizes R_1 and R_2 and velocities of the bodies. This is usually expressed in form of the gravitational focusing factor F_g [48, 68]:

$$F_g = 1 + \frac{v_{\mathrm{esc}(1,2)}^2}{\sigma^2}, \tag{2.2}$$

in the collisional cross section Γ

$$\Gamma = \pi R_s^2 F_g, \tag{2.3}$$

where

$$v_{\mathrm{esc}(1,2)} = \sqrt{\frac{2G(m_1 + m_2)}{R_s}} \tag{2.4}$$

is the escape velocity from the point of contact, $R_s = R_1 + R_2$ is the sum of the physical radii and σ represents the relative velocities of the bodies at infinity. If some larger bodies of mass m_1 are situated in a swarm of smaller particles of mass $m_2 \ll m_1$, the overall dynamics is governed by the small bodies and σ is of the order of $v_{\mathrm{esc}(2,2)}$. F_g for the smaller particles is of order unity, whereas F_g for the larger particles is much larger, because $v_{\mathrm{esc}(2,2)} \ll v_{\mathrm{esc}(1,2)}$. Hence, at this stage, there is not any orderly continuous growth of all planetesimals of different sizes. In fact, the largest planetesimals grow very quickly and detach from the initial size distribution [47]. This is mainly caused by the decrease of random velocities with increasing planetesimal mass [69].

The growing bodies are often called *planetary embryos*. While the planetary embryos grow, they gravitationally perturb nearby planetesimals more and more and their own velocities are kept small by gravitational friction. Hence, the small bodies are not able to enter runaway growth. They are disrupted via high-energy mutual collisions or accreted by a planetary embryo and participate in their growth. Simultaneously, the growth rate of the planetary embryos decreases [70]. This phase,

where the big ones get bigger and the small ones get nothing is called the *oligarchic growth* [71].

The ultimate end of runaway growth is given when the planetary embryos have accreted all planetesimals within their gravitational reach [72]. The final mass of such a planetary embryo, called the isolation mass $M_{\rm iso}$, is roughly given by the initial mass inside an annulus of half width of several Hill-radii $\Delta r = B\, R_{\rm Hill}$, called the *feeding zone* or *source zone*. The Hill-radius $R_{\rm Hill}$ of a body at semi-major axis r of mass M_p orbiting a star of mass $M_\star$ is

$$R_{\rm Hill} \equiv \left(\frac{M_p}{3M_\star}\right)^{1/3} r. \tag{2.5}$$

The isolation mass at semi-major axis r is then given by:

$$M \approx 4\pi r \Delta r \Sigma = \frac{(4\pi B r^2 \Sigma)^{3/2}}{(3M_\star)^{1/2}}, \tag{2.6}$$

where Σ is the local surface density of planetesimals.

Thus, with $B = 2\sqrt{3}$, the feeding zone of a planetary embryo at 1 AU has a typical width of some percent of an AU and the $M_{\rm iso} \approx 0.07\, M_\oplus$ is around a few times the mass of the Moon [47]. In contrary, at 5 AU, the isolation mass can reach $M_{\rm iso} \approx 9\, M_\oplus$ [73], which provides a fundament to form the cores of the giant planets. The masses are just rough estimates, as the planetesimals can diffuse into the accretion zone of a planetary embryo via scattering or gas drag [74] and barely all planetesimals in the feeding zone are accreted before entering the next stage. Nevertheless, this model claims that the formation of planetary embryos is constrained to relatively small annuli in the protoplanetary disk. Radial mixing, the exchange of material over large distances, has not yet taken place.

The major drawback in this picture of embryo formation is the collisional growth. Recent studies have shown that the collision between planetesimals are often disruptive since the random velocities of the planetesimals are increased by turbulences in the gas disk (e.g. [75]) and that planetesimals tend to fragment faster than expected [76]. Hence, the collisional growth starting with planetesimals has to be bypassed, which is possible if some bodies with a diameter of 100–1000 km are available at the beginning of this formation stage. This is provided when planetesimals form via e.g. streaming instabilities, as mentioned above.

2.2.4. From planetary embryos to terrestrial planets

Compared to the initial stages, which take about 0.1–1 Myr, the last stage is the slowest. It can take 10–100 Myr or more, depending on the system architecture and on the disk properties. In the final stage of terrestrial planet formation, the planetary embryos start to perturb one another. As long as the gas disk is present with significant surface density, it damps continuously the eccentricities of embedded bodies. The same holds for damping thanks to dynamical friction induced by the remaining planetesimals. When the gas and the majority of the planetesimals are removed, the eccentricities of the embryos' orbits start to grow rapidly, because the

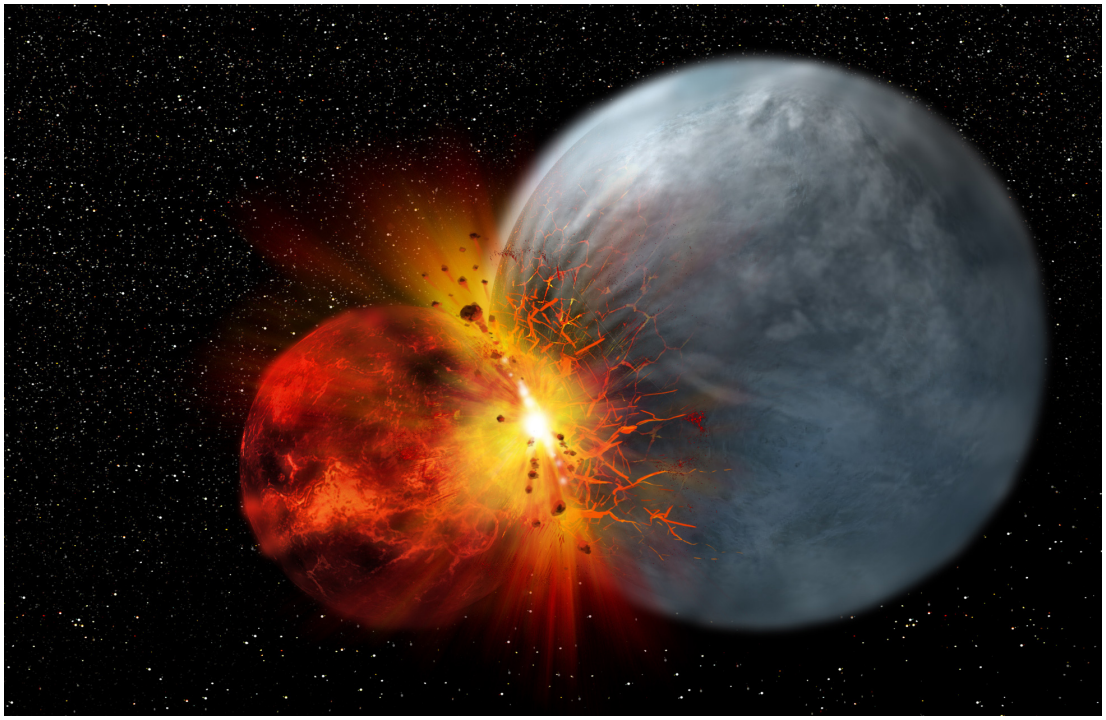

Figure 2.3. Artist's impression of a giant impact. Such impacts are the main mechanism of growth in mass in the final stage of planet formation. In addition, they initiate the formation of massive satellites like the Moon. (credit: Michael Elser)

embryos are not separated enough to stay on constant orbits over a long period of time. This excitement of the orbital eccentricities results in intersecting orbits, close encounters and collisions and finally in the formation of terrestrial planets. Perturbations by giant planets in the system can increase the eccentricities even more and reduce the accretion time scale [77]. We refer to this phase as the *giant impact* phase.

This stage is responsible for many physical and chemical characteristics of the final planets. The mass and orbital parameters like semi-major axis or eccentricity are strongly dependent on properties of the planetary embryos that form the planet. Although remaining planetesimals can still take part in the growth of the embryos, giant impacts add most of the mass [78]. Beside external forces like resonances with the other planets or migration in the remnants of the gas disk, the collisions have a strong effect on the final orbit of the planet.

The stochastic sequence of giant impacts is also responsible for the final spin of the planets. Most of a prograte planet's spin is originated by the last major impact [79]. However, the accretion of small planetesimals could result in a systematic component in the rotation state of a planet [80].

Planetary embryos are pushed to higher eccentricities and can be scattered, which results in a significant change in semi-major axis. In contrary to the runaway growth phase, where only material from a narrow feeding zone is comprised in the planetary embryos, the giant impact phase provides strong radial mixing. A detailed study on the effect of this kind of radial mixing on the final elemental abundances in terrestrial

planet is shown in chapter 4.

When planetary embryos collide (figure 2.3), the collisions are far from being simple perfect mergers. Depending on impact angle, velocity and mass of the embryos, a large amount of material can be transported into a circumplanetary orbit and lays the basis for satellite formation [81–83]. Collisions suitable for impact-generated satellite formation are not rare, see 3 for details. Giant impact can also fundamentally change the bulk abundances of planetary bodies. Impact release a large amount of energy and large parts of the bodies can heat up to several 1000 K. This can result in the loss of volatile elements initially comprised in the bodies [84]. A gracing impact might be responsible for the stripping off of Mercury's silicate mantle [85], resulting in a large bulk abundance of heavy elements like iron. Fragments and unbound debris can result in dynamical friction which may play an important role at this stage by damping orbital elements of the embryos or final planets [86, 87].

2.2.5. Formation of giant planets

Compared to terrestrial planets, giant planets form quickly. Currently, there are two established models that explain giant planet formation. First, we briefly present the *core accretion* model [73, 88]. The formation process starts similar to terrestrial planet formation, beyond the snow line, where water and other ices can condense. As we have seen, planetary embryos that form in this region can grow up to $10\,M_\oplus$ in a time scale of 1 Myr. During its growth, it becomes massive enough to accrete an envelope of hydrogen and helium. At first, this envelope is in hydrostatic equilibrium and grows with further accretion of gas. Planetesimals are accreted at the same time and provide the last part of the mass of the core. The energy released when the planetesimals impact on the core counteracts the collapse of the envelope. If the core reaches a critical mass, gravity dominates and the gas envelope contracts and rapid gas accretion takes place. This final accretion phase stops when the gas planet opens a gap in the disk, thus cuts off the supply of gas, or when the gas disk is dispersed. Hence, a key aspect in this model is that the core formation is quickly enough to provide substantial gas accretion.

The other model for giant planet formation is the *gravitational instability* model [89–91]. In this model, the gas disk becomes gravitational unstable and fragments into bound, self-gravitating clumps. In contrary to the core accretion model, planets form directly out of gas within a very short period of time. Then, the core is formed via sedimentation of solids and accretion of planetesimals. The conditions for disk fragmentation depend on the surface density, sound speed and distance from the star [61]. A sufficient high surface density might be given at around 10 AU in a very young protoplanetary disk, which is much more massive than the MMSN [48]. The sound speed is controlled by the temperature in the disk and the latter depend on the disk's ability to cool [92]. Numerical simulations show that disk fragmentation is possible at large radii > 10 AU [93, 94]. Some simulations show fragmentation also at smaller radii [95].

Given by the nature of their composition and their short formation time scale, giant planets are formed completely when the gas disk disappears. Hence, the architecture and early evolution of the giant planets in a system are of fundamental

importance in the formation process of the terrestrial planets. In general, the orbital properties, mass and number of terrestrial planets that form depend on the orbital properties, mass and number of the giant planets in the system. Excitation of the planetary embryos through e.g. secular resonances result in more collisions, which results in a small number of large planets [96, 97]. On the other hand, the absence of giant planets results in a lack of dynamical excitation and the terrestrial planet formation proceeds on a longer time scale.

2.2.6. Dynamical evolution of young planetary systems

During and after the formation of a planetary systems, the orbits of the planets can evolve significantly due to interaction with other planets or with the remnants of the formation process.

As long as the gas disk is present, planets (or planetary embryos) interact with the gas disk as a consequence of energy and angular momentum exchange between the planet and density waves that are caused by the planet in the gas disk. Low mass planet do not change the disk structure fundamentally. Still, spiral density waves can act a torque on the planet and move it inwards or outwards. This effect is called type 1 migration. Higher mass planets are able to open an annular gap in the gas disk and migrate then together with this gap. This is called type 2 migration. The interaction of remnant planetesimals results in planetesimal-driven migration, which might even dominate gas-driven migration under certain conditions [98]. Orbital migration of gas giants is a possible explanation for the existence of hot Jupiters. It can also explain some features in the Solar System. For example, the small mass of Mars can be explained by the inward migration of Saturn and Jupiter until they get in resonance around 1.5 AU and migrate outwards. This violent intrusion in the planetesimal belt inside the snow line truncated the feeding zone of the outer planetary embryos significantly and results in a small mass of the outermost terrestrial planet [99]. Yet another example is the given by the Nice Model [100], which is based on the planetesimal-driven orbital movement of the early giant planets to their final location. It provides a model for the origin of the Late Heavy Bombardment, the Oort cloud and other features in the Solar System.

Of course, radial migration can result in violent mutual interaction between planets. In the Nice model, planet-planet scattering results in a interim high eccentricity of Neptune. It can start to clear the outer remnants of the planetesimal disk before its eccentricities is damped again by dynamical friction.

In chapter 5, we study the long-term stability of known planetary systems under the assumption that hypothetical Super-Earth mass planets are present. This enables us to predict the most probable orbital elements of potential Super-Earths in the HZ. In return, this work shows that planetary systems can form in a configuration which is not dynamically stable over a main-sequence life time. When instabilities arise, planet-planet scattering can even result in the ejection of planets and in high eccentricities of the remaining ones. Such scattering events can explains the high average eccentricities observed in extrasolar systems, as mentioned before.

2.3. Simulating the formation and evolution of planets

N-body simulations provide a laboratory to test models of planet formation and the dependence of the outcome on initial conditions. Runaway growth and especially the giant impact phase are suitable to be studied with such simulations as well as the long-term evolution of planetary systems. We briefly give some insights in the numerical methods and show how successful present day simulations are.

2.3.1. Numerical methods

A system of N particles is described by the positions $\mathbf{r}$, velocities $\mathbf{v}$ and masses m of the particles. The time evolution of such a system is governed by $6N$ coupled differential equations of first order. The motion of particle i is given by:

$$\frac{\mathrm{d}\mathbf{r_i}}{\mathrm{d}t} = \mathbf{v_i}, \quad \frac{\mathrm{d}\mathbf{v_i}}{\mathrm{d}t} = \frac{\mathbf{F_i}}{m_i}. \tag{2.7}$$

where $\mathbf{F_i}$ denotes the vector sum of all forces acting on the particle i. As we have seen, there are several forces that act on a body during the late stages of planet formation: gas drag, torques from density waves in the gas disk, the gravity of the gas disk, planetesimals, planetary embryos, planets and the central star. The calculations of the forces are in principle straightforward. In fact, due to the high number of small bodies like planetesimals and planetary embryos, they are the most time consuming computations.

In the following, we want to briefly describe two codes which allow the simulation of the orbital movement and interaction of thousands of bodies over hundreds of Myr. These codes use different methods when calculating the forces between the bodies. The first code, GENGA, was used in chapter 5 to integrate the orbital motion of planetary systems, the second code, pkdgrav planet, carried out the planet formation simulations used in chapters 3 and 4.

Direct integration - GENGA

The GENGA Code is a hybrid symplectic integrator [101], based on the widely used Mercury code [102]. It consists of two different integration methods. As long as the interaction between two bodies are small, the gravitational forces between bodies are computed as perturbations of Keplerian orbits using a mixed variable integrator [103]. This integrator breaks down when two bodies come close and their mutual perturbations become significant compared to the Keplerian motion. In this case, the code switches smoothly to a direct N-body Bulirsch-Stoer integrator, which integrates (2.7) up to high precision during a close encounter. Using this serial method, the computing time per time step scales basically with N^2. Nevertheless, since GENGA runs on graphics processing units (GPUs), it can perform many computations in parallel and has an advantage over similar serial CPU codes.

Tree method - pkdgrav planet

In a tree method like PKDGRAV, the force calculations are made by dividing all particles into hierarchical tree-structured groups [104]. Then, the gravitational forces

are given by summing up the individual contribution from each group based on their multipole expansion. In Morishima et al. (2010) [78], this method is adapted to planet formation studies. The PKDGRAV PLANET code is base on the SyMBA mixed symplectic integrator, while the tree code provides a fast calculation of the perturbations of the orbits. Compared to the direct summation in GENGA, the tree method is a bit less exact, since the forces of distant particles are grouped. Similar to GENGA, close encounters are separately calculated with high precision. Since only a small fraction of all particles is in a close encounter during one time step, the tree code does most of the force calculations. This results in a total computation time per time step that scales as $N \cdot logN$. Hence, the tree method is faster than serial methods when the number of particles is high. Since the tree building process takes some time, it can only unfold its power when $N \gtrsim 1000$.

2.3.2. Present day simulations

Numerous numerical experiments have been carried out during the last three decades [77, 78, 97, 105–110]. They provide insight into the dependence of the resulting terrestrial planets on the simulation and model parameters. Given by the fact that until the discovery of the first exoplanets in the past years, the inner Solar System was the only known configuration of rocky planets, most simulations and models are designed to reproduce the Solar System planets. Hence, results can be compared with constraints given by the Solar System and the simulations and models can be improved and expanded. Some simulations are explicitly carried out to study terrestrial planet formation in other planetary systems and thus, other configurations of large planets are involved, e.g. [111]. As mentioned several times, the architecture of the Solar System might not be common. In the future, when more terrestrial planets are known, simulations taking into account the huge variety of planetary system configurations will become important and groundbreaking.

The typical initial configuration in present day Solar System formation simulations is given as follows. The fundament is a gas disk. Accounting for subsequent removal of the gas, the surface density changes as a function of time and/or space. A belt of planetesimals is embedded in the gas disk. These bodies can be of equal size, they can follow a size distribution or the belt can comprise of particles of different size regimes (e.g. planetesimals and planetary bodies). Since secular perturbation through the gas giants is of great importance, often the fully formed Jupiter and Saturn are included. Their orbits can be similar to the present day orbits or more circular and with a smaller separation (according to the *Nice* model [100]), accounting for the further evolution of the giants' orbits.

A descriptive example is shown in figure 2.4. It presents some snapshots of a simulation carried out by Morishima et al. (2010) [78], whose simulations are used extensively in chapters 3 and 4. Initially, a disk of equally sized planetesimals with total mass of $10 M_\oplus$ is distributed according to a r^{-1} power-law in surface density. They are embedded in a gas disk following the same radial power-law which is decaying exponentially on a 1 Myr time scale. Accordingly, giant planet formation is limited roughly to this time span and the giant planets Jupiter and Saturn are introduced on circular orbits after 1 Myr. Secular resonances with the giants and

scattering among the planetesimals increases the eccentricities of the bodies, which results in a higher collision rate and a faster growth of the particles. After 3 Myr, the inner part of the belt is dominated by planetary embryos which cleaned their feeding zone from planetesimals. Planetesimals and embryos can be pushed to eccentric orbits when the damping of the gas disk has vanished almost completely. Giant impacts take place and the formation of massive satellites can be induced. The bottom panel of figure 2.4 shows how the planets in the simulation grow via accretion of small bodies and planetary embryos. The giant impact phase starts after approximately 1 Myr.

Simulations are rather successful in reproducing features of the Solar System. Two to four planets on stable orbits are typically formed inside 0.5–2 AU. Depending on the initial disk mass and the orbits of the giant planets, the mass of the largest planet can be close to 1 $M_\oplus$. N-body simulations including only the interaction of planetary bodies result often in the overestimation of the final eccentricity and inclination [77, 107, 108]. It is therefore indispensable to include a population of smaller bodies in embyro based simulations or to start directly with a distribution of planetesimal-sized bodies. Then, damping by dynamical friction of the remnant planetesimals results in eccentricities and inclinations comparable with the Solar System [110, 112]. The gas disk affects also the planetesimal orbits fundamentally [113] and should be taken into account [78]. Morishima et al. (2010) pointed out the trade-off of the resulting mass distribution of the planets and the timing of the Moon-forming impact. In addition, all of these simulations usually fail in reproducing the small mass of Mars. As mentioned above, a solution to this problem was provided recently [99]: assuming that Jupiter and Saturn migrated inwards (and then outwards) after their formation, the disk of planetesimals and planetary embryos was truncated and only a very low mass planet could form at the location of Mars.

Some problems of present day simulations originate from the simple formation model and from the limited possibilities to study the dependence on initial and boundary conditions. The mutual collisions in the N-body simulations are usually treated as perfect mergers: when bodies collide, their momentum and their mass are conserved and a new body is formed. This is truely an oversimplification. Nevertheless, its effect on planet formation simulations seems to be small [114]. Of course, the inclusion of fragmentation in the N-body simulations would be desirable, since it could provide direct insights in potential satellite-forming impacts and in the effect of dynamical friction on the planetary embryos due to collisional generated remnants. We discuss further implications in chapter 6.

The dependence on initial and boundary conditions is studied in many simulations. The initial mass, profile and limits of the gas and planetesimals disk, the way the gas disk disappears, the orbits of Jupiter and Saturn and their evolution: all these variables span a huge multi-dimensional parameter space which is only partially study yet. This holds for the Solar System, but other planetary systems show completely different configurations.

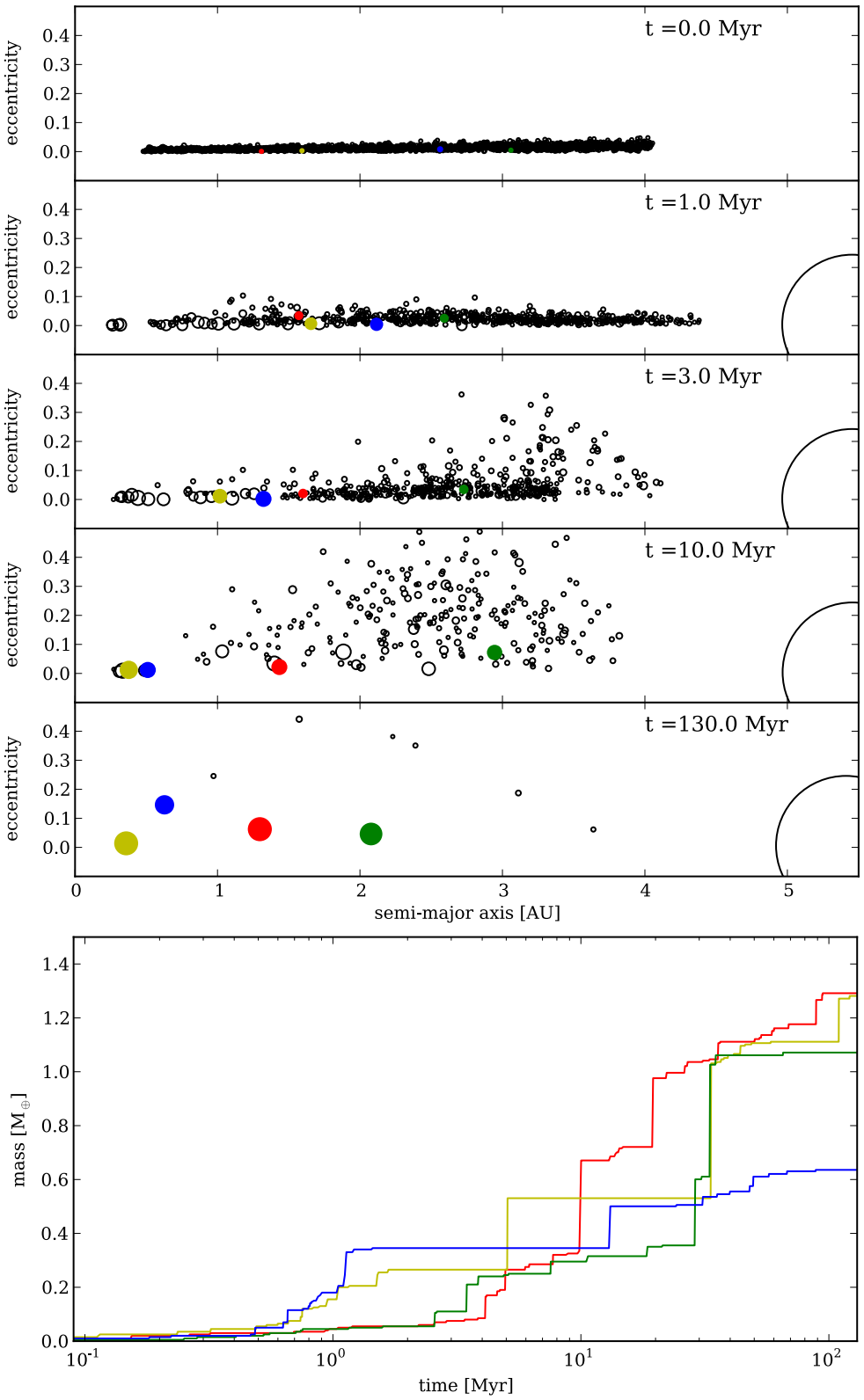

Figure 2.4. The growth of terrestrial planets in a N-body simulation by Morishima et al. (2010) [78]. The top panel shows snapshots of the formation in the (a,e)-plane. The bottom panel shows the growth of the planets with time. Each color is assigned to a planet.

3

Paper I
Earth-Moon Systems

In this chapter, we estimated the occurrence rate of Earth-Moon planetary systems. The formation of a massive satellite like the Moon by itself is a fascinating event worth to be studied in detail. Nevertheless, we were mainly motivated by the importance of the Moon concerning the emergence and survival of life on the Earth. The early Moon was much closer to our planet than today causing high tides several times per day. This could support the formation of early biomolecules in coastal regions. Furthermore, the massive Moon stabilizes the spin axis of our planet and prevent it from chaotic motion, which clearly plays to a stable climate. This important role of the Moon could also hold for Earth-like planets in extrasolar system. It is a small piece in the complex notion of habitability.

We found that Earth-Moon binary systems are not rare. 1 in 12 Earth-like planets host a massive satellite. This is a relatively high fraction, which indicates that many Earth-like planets can profit from the positive effect of hosting a massive satellite.

This paper was published 2011 in Icarus [115].

Icarus 214 (2011) 357–365

Contents lists available at ScienceDirect

Icarus

journal homepage: www.elsevier.com/locate/icarus

How common are Earth–Moon planetary systems?

S. Elser [a], B. Moore [a,*], J. Stadel [a], R. Morishima [b]

[a] University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[b] LASP, University of Colorado, Boulder, CO 80303-7814, USA

ARTICLE INFO

Article history:
Received 5 November 2010
Revised 17 May 2011
Accepted 18 May 2011
Available online 22 June 2011

Keywords:
Moon
Terrestrial planets
Planetary formation
Satellites, Formation

ABSTRACT

The Earth's comparatively massive moon, formed via a giant impact on the proto-Earth, has played an important role in the development of life on our planet, both in the history and strength of the ocean tides and in stabilizing the chaotic spin of our planet. Here we show that massive moons orbiting terrestrial planets are not rare. A large set of simulations by Morishima et al. (Morishima, R., Stadel, J., Moore, B. [2010]. Icarus. 207, 517–535), where Earth-like planets in the habitable zone form, provides the raw simulation data for our study. We use limits on the collision parameters that may guarantee the formation of a circumplanetary disk after a protoplanet collision that could form a satellite and study the collision history and the long term evolution of the satellites qualitatively. In addition, we estimate and quantify the uncertainties in each step of our study. We find that giant impacts with the required energy and orbital parameters for producing a binary planetary system do occur with more than 1 in 12 terrestrial planets hosting a massive moon, with a low-end estimate of 1 in 45 and a high-end estimate of 1 in 4.

1. Introduction

The evolution and survival of life on a terrestrial planet requires several conditions. A planet orbiting the central star in its habitable zone provides the temperature suitable for the existence of liquid water on the surface of the planet. In addition, a stable climate on timescales of more than a billion years may be essential to guarantee a suitable environment for life, particularly land-based life. Global climate is mostly influenced by the distribution of solar insolation (Milankovitch, 1941; Berger et al., 1984; Berger, 1989; Atobe and Ida, 2006). The annual-averaged insolation on the surface at a given latitude is, beside the distance to the star, strongly related to the tilt of the rotation axis of the planet relative to the normal of its orbit around the star, the obliquity. If the obliquity is close to 0°, the poles become very cold due to negligible insolation and the direction of the heat flow is poleward. With increasing obliquity, the poles get more and more insolation during half of a year while the equatorial region becomes colder twice a year. If the obliquity is larger than 57°, the poles get more annual insolation than the equator and the heat flow changes. Therefore, the equatorial region can even be covered by seasonal ice (Ward and Brownlee, 2000). Thus, the obliquity has a strong influence on a planet's climate. The long-term evolution of the Earth's obliquity and the obliquity of the other terrestrial planets in the Solar System, or planets in general, is controlled by spin–orbit resonances and the tidal dissipation due to the host star and satellites of the planet. Thus, the evolution of the planetary obliquity is unique for each planet.

Earth's obliquity fluctuates currently ±1.3° around 23.3° with a period of ~41,000 years (Laskar et al., 1993; Laskar and Robutel, 1993; Laskar, 1996). The existence of a massive (or close) satellite results in a higher precession frequency which avoids a spin–orbit resonance. Without the Moon, the obliquity of the Earth would suffer very large chaotic variations. The other terrestrial planets in the Solar System have no massive satellites. Venus has a retrograde spin direction, whereas a possibly initial more prograde spin may have been influenced strongly by spin–orbit resonances and tidal effects (Goldreich and Peale, 1970; Laskar, 1996). Mars' obliquity oscillates ±10° degree around 25° with a period of several 100,000 years (Ward, 1974; Ward and Rudy, 1991). Mercury on the other hand is so close to the Sun that its rotation period is in an exact 3:2 resonance with its orbital period. Mercury's spin axis is aligned with its orbit normal.

On larger timescales, the variation of the obliquity can be even more dramatic. It has been shown that the tilt of Mars' rotation axis ranges from 0° to 60° in less than 50 Myr and 0° to 85° in the case of the obliquity of an Earth without the Moon (Laskar et al., 1993; Laskar and Robutel, 1993; Laskar, 1996).

The main purpose of this report is to explore the giant impact history of the planets in order to calculate the probability of having a giant Moon-like satellite companion, based on simulations done by Morishima et al. (2010). A giant impact between a planetary embryo called *Theia*, the Greek titan that gave birth to the Moon goddess *Selene*, first named by Halliday (2000), and the proto-Earth is the accepted model for the origin of our Moon (Hartmann and

* Corresponding author.
E-mail addresses: selser@physik.uzh.ch (S. Elser), moore@physik.uzh.ch (B. Moore), stadel@physik.uzh.ch (J. Stadel), ryuji.morishima@lasp.colorado.edu (R. Morishima).

0019-1035/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.icarus.2011.05.025

Davis, 1975; Cameron and Ward, 1976; Cameron and Benz, 1991), an event which took place within about 100 Myr after the formation of calcium aluminum-rich inclusions in chondritic meteoroids, the oldest dated material in the Solar System (Touboul et al., 2007). After its formation, the Moon was much closer and the Earth was rotating more rapidly. The large initial tidal forces created high tidal waves several times per day, possibly promoting the cyclic replication of early bio-molecules (Lathe, 2004) and profoundly affecting the early evolution of life. Tidal energy dissipation has caused the Moon to slowly drift into its current position, but its exact orbital evolution is still part of an on-going debate (Varga et al., 2006; Lathe, 2006). Calculating the probability of life in the Universe (Ward and Brownlee, 2000) as well as the search for life around nearby planets may take into account the likelihood of having a massive companion satellite.

This report is structured as follows: In Section 2, we give a brief review on the evolution of simulating terrestrial planet formation with N-body codes during the last decades and present the method we used. In Section 3, we study the different parameters of a protoplanet collision to identify potential satellite forming events. In Section 4, we summarize the different uncertainties from the simulations and our analysis that may affect the final results. Finally, we give a conclusion, we present our results and compare them with previous works in Section 5.

2. Simulating terrestrial planet formation

There are good observational data on extra-solar gas giant planets, but whilst statistics on extra-solar rocky planets will be gathered in the coming years, for constraints on the formation of the terrestrial planets we rely on our own Solar System. The established scenario for the formation of the Earth and other rocky planets is that most of their masses were built up through the gravitational collisions and interactions of smaller bodies (Chamberlin, 1905; Safronov, 1969; Lissauer, 1993). Wetherill and Stewart (1989) observed the phase of run-away growth. This phase is characterized by the rapid growth of the largest bodies. While their mass increases, their gravitational cross section increases due to gravitational focusing. When a body reaches a certain mass, the velocities of close planetesimals are enhanced, the gravitational focusing decreases and so does the accretion efficiency. This is called the oligarchic growth phase, first described by Kokubo and Ida (1998). During this phase, the smaller embryos will grow faster than the larger ones. At the end, several bodies of comparable size are embedded in a planetesimal disk. These protoplanets merge via giant impacts to form the final planets. Dones and Tremaine (1993) showed that most of a terrestrial planet's prograde spin is imparted by the last major impactor and can not be accumulated via the ordered accretion of small planetesimals. Giant impacts with a certain impact angle and velocity generate a disk of ejected material around the target which is a preliminary step in the formation of a satellite. Usually, the simulations assume perfect accretion in a collision. Tables of the collision outcome can help to improve the simulations or to estimate the errors in the planetary spin, (Kokubo and Genda, 2010). Until recently simulations were limited in the number of planetesimal bodies that could be self-consistently followed for time spans of up to billions of years, but recent algorithmic improvements by Duncan et al. (1998) in his SyMBA code and by Chambers (1999) in his Mercury code have allowed them to follow over long time spans a relatively large number of bodies ($\mathcal{O}(1000)$) with high precision, particularly during close encounters and mergers between the bodies, where individual orbits must be carefully integrated (Chambers and Wetherill, 1998; Agnor et al., 1999; Raymond et al., 2004; Kokubo et al., 2006). Raymond et al. (2009) have also recently conducted a series of simulations where they varied the initial conditions for the gas giant planets and also track the accretion of volatile-rich bodies from the outer asteroid belt, leading either to "dust bowl" terrestrial planets or "water worlds" and everything between these extremes. All prior simulation methods with full interaction among all particles have however been limited in number of particles since their force calculations scale as $\mathcal{O}(N^2)$.

We have developed a new parallel gravity code that can follow the collisional growth of planetesimals and the subsequent long-term evolution and stability of the resulting planetary system. The simulation code is based on an $\mathcal{O}(N)$ fast multipole method to calculate the mutual gravitational interactions, while at the same time following nearby particles with a highly accurate mixed variable symplectic integrator, which is similar to the SyMBA (Duncan et al., 1998) algorithm. Since this is completely integrated into the parallel code PKDGRAV2 (Stadel, 2001), a large speed-up from parallel computation can also be achieved. We detect collisions self-consistently and also model all possible effects of gas in a laminar disk: aerodynamic gas drag, disk-planet interaction including Type-I migration, and the global disk potential which causes inward migration of secular resonances with gas dissipation. In contrast to previously mentioned studies, this code allows us to self-consistently integrate through the last two phases of planet formation with the same numerical method while using a large number of particles.

Using this new simulation code we have carried out 64 simulations which explore sensitivity to the initial conditions, including the timescale for the dissipation of the solar nebula, the initial mass and radial distribution of planetesimals and the orbits of Jupiter and Saturn (Morishima et al., 2010). All simulations start with 2000 equal-mass particles placed between 0.5 and 4 AU. The initial mass of the planetesimal disk m_d is 5 or $10m_\oplus$. The surface density Σ of this disk and of the initial gas disk depends on the radius through $\Sigma \propto r^{-p}$, where p is 1 or 2. The gas disk dissipates exponentially in time and uniformly in space with a gas dissipation time scale τ_{gas} = 1, 2, 3 or 5 Myr. After the disappearance of the gas disk (more precisely after time τ_{gas} from the beginning of the simulation), Jupiter and Saturn are introduced on their orbits, e.g. circular orbits or the current orbits with higher eccentricities.

Fig. 1 shows two merger trees: a $1.8m_\oplus$ planet formed in the simulation with (τ_{gas}, p, m_d) = (1 Myr, 1, $5m_\oplus$) and gas giants on the present orbits and a 1.1 $m_\oplus$ planet formed in the simulation with (τ_{gas}, p, m_d) = (1 Myr, 2, $10m_\oplus$) and gas giants on circular orbits. The red branches are satellite forming impactors both with a mass $0.3m_\oplus$ and represent two events of the final sample in Fig. 7. These merger trees with their different morphologies reveal the variety of collision sequences in terrestrial planet formation. They show that a large set of impact histories is generated by these simulations despite the relatively narrow parameter space for the initial conditions.

3. Satellite formation

During the last phase of terrestrial planet formation, the giant impact phase, satellites form. Collisions between planetary embryos deposit a large amount of energy into the colliding bodies and large parts of them heat up to several 10^3 K, e.g. Canup (2004). Depending on the impact angle and velocity and the involved masses, hot molten material from the target and impactor can be ejected into an circumplanetary orbit. This forms a disk of ejecta, the disk material is in a partially vapor or partially molten state, around the target planet. The proto-satellite disk cools and solidifies. Solid debris form and subsequently agglomerate into a satellite (Ohtsuki, 1993; Canup and Esposito, 1996; Kokubo et al., 2000).

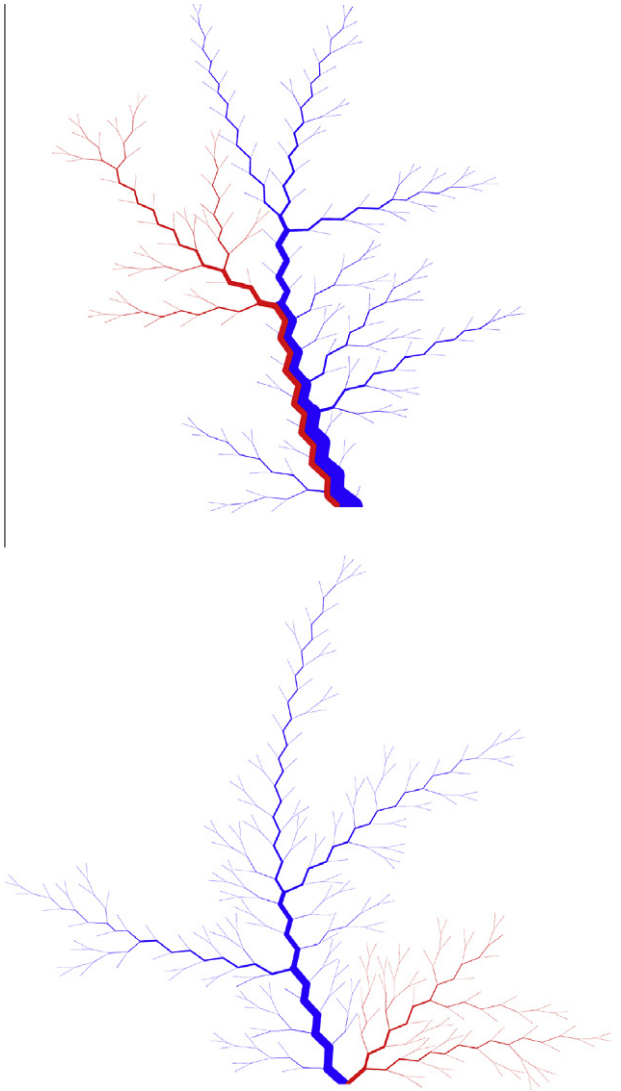

Fig. 1. Two merger trees. They illustrate the accretion from the initial planetesimals to the last major impactors that merge with the planet. Every 'knee' is a collision of two particles and the length between two collisions is given by the logarithm of the time between impacts. The thickness of the lines indicates the mass of the particle (linear scale). The red branch is the identified satellite forming impact in the planet's accretion history. *Top:* a $1.1m_{\oplus}$ planet formed in the simulation with (τ_{gas}, p, m_d) = (1 Myr, 2, $10m_{\oplus}$) and gas giants on circular orbits and a $0.3m_{\oplus}$ impactor. In this case, the moon forming impact is not the last collision event but it is followed by some major impacts. *Right:* a $0.7m_{\oplus}$ planet formed in the simulation with (τ_{gas}, p, m_d) = (3 Myr, 1, $5m_{\oplus}$) and gas giants on the present orbits and a $0.2m_{\oplus}$ impactor. It is easy to see that the moon forming impact is the last major impact on the planet. Although this planet is smaller than the upper one, it is composed out of a similar number of particles. In the case of the more massive disk (m_d = 10), the initial planetesimals are more massive since their number is constant. Therefore, fewer particles are needed to form a planet of comparable mass than in the case of m_d = 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The giant impact which resulted in the Earth–Moon system is a very particular event (Cameron and Benz, 1991; Canup, 2004). The collision parameter space that describes a giant impact can by parametrized by $\gamma = m_i/m_{tot}$, the ratio of impactor mass m_i to total mass in the collision m_{tot}, by $v \equiv v_{imp}/v_{esc}$, the impact velocity in units of the escape velocity $v_{esc} = \sqrt{2Gm_{tot}/(r_i + r_t)}$, where r_t and r_i are the radii of target and impactor. Furthermore, it is described by the scaled impact parameter b, where $b = 0$ indicates a head-on collision and $b = 1$ a grazing encounter, and the total angular momentum L. Recent numerical results (Canup, 2008) obtained with smoothed particle hydrodynamic (SPH) simulations for the Moon-forming impact parameters require: $\gamma \sim 0.11$, $v \sim 1.1$, $b \sim 0.7$ and $L \sim 1.1L_{EM}$, where L_{EM} is the angular momentum of the present Earth–Moon system. These simulations also include the effect of the initial spins of the colliding bodies, but the explored parameter space is restricted to being close to the Moon-forming values given above.

If one does not focus on a strongly constrained system like the Earth–Moon system but just on terrestrial planets of arbitrary mass with satellites that tend to stabilize their spin axis, the parameter space is broadened. It becomes difficult to draw strict limits on the parameters because collision simulations for a wider range of impacts were not available for our study. Hence, based on published Moon-forming SPH simulations by Canup (2004, 2008), we use a semi-analytic expression to constrain the mass of a circumplanetary disk that can form a satellite. In addition, we include tidal evolution and study the ability of the satellite to stabilize the spin axis of the planet.

3.1. Satellite mass and collision parameters

We do not know the exact outcome of a protoplanet collision, but certainly the satellite mass is related to the collision parameters of the giant impact. Hence, we can draw a connection from these parameters to the mass of the final satellite. Based on the studies of the Earth's Moon formation, we can start with a simple scaling relation: a Mars-size impactor gives birth to a Moon-size satellite. Their mass ratio is $m_{Mars}/m_{Moon} \sim 10$. Thus, to very first approximation, we can assume that an impactor mass is usually 10 times larger than the final mass of the satellite. Of course, this shows that only a small amount of material ends up in a satellite, but this statement is only valid for a certain combination of mass ratio, impact parameter and impact speed and usually gives an upper limit on the satellite mass.

In order to get a better estimation of the satellite mass, we use the method obtained in the appendix of Canup (2008).

There, an expression is derived that describes the mass of the material that enters the orbit around the target after a giant impact:

$$\frac{m_{disk}}{m_{tot}} \sim C_{\gamma}\left(\frac{m_{pass}}{m_{tot}}\right)^2, \tag{1}$$

where the prefactor $C_{\gamma} \sim 2.8(0.1/\gamma)^{1.25}$ has been determined empirically from the SPH data. m_{pass}/m_{tot} is mass of the impactor that avoids direct collision with the target. It depends mainly on the impact parameter b and on the mass ratio γ and can be computed by studying the geometry of the collision. The total impactor volume that collides with the target is:

$$V_T = \int_0^{\pi} A(\phi)d\phi, \tag{2}$$

with

$$A(\phi) = r_t^2\theta_t(\phi) + r_i^2\theta_i(\phi) - Dr_i\sin\phi\sin\theta_i(\phi), \tag{3}$$

where r_i and r_t are impactor and target radius, $D = b(r_i + r_t)$ gives the distance between the centers of the bodies and

$$\theta_i(\phi) = \cos^{-1}\left[\frac{D^2 + r_i^2\sin^2\phi - r_t^2}{2Dr_i\sin\phi}\right], \tag{4}$$

$$\theta_t(\phi) = \cos^{-1}\left[\frac{D^2 + r_t^2 - r_i^2\sin^2\phi}{2Dr_t}\right]. \tag{5}$$

Assuming a differentiated impactor with $r_{core} \sim 0.5r_i$ and repeating the above integration for this radius, the colliding volume of the impactor mantle is $V_{mantle} = V_T - V_{core}$. If we assume

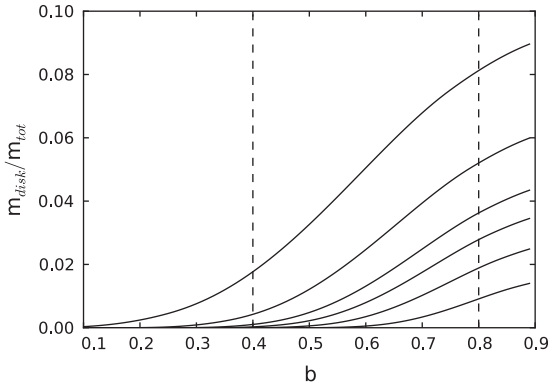

Fig. 2. The disk mass resulting after a giant impact in units of the total mass of the colliding system relative to the impact parameter b based on Eq. (1). The different solid lines belong to different mass ratios γ. From bottom to top: $\gamma = 0.05$, 0.1, 0.15, 0.2, 0.3 and 0.5. These mass ratios result in different disk masses. The equation is valid up to a factor 2 in between the two dash lines ($0.4 < b < 0.8$) for small velocities. The disk mass is an upper bound on the satellite mass.

that the core is iron and the mantle dunite, the core density ρ_{core} has roughly twice the density of the mantle ρ_{mantle}. The mass of the impactor that hits the target is $m_{\text{hit}} = \rho_{\text{core}} V_{\text{core}} + \rho_{\text{mantle}} V_{\text{mantle}}$. Therefore, the mass that passes the target is $m_{\text{pass}} = m_i - m_{\text{hit}}$.

We use the full expression derived by Canup (2008), Eq. (1), to estimate the disk mass resulting from the giant impacts in our simulation. Eq. (1) is correct to within a factor 2, if $v < 1.4$ and $0.4 < b < 0.7$ or if $v < 1.1$ and $0.4 < b < 0.8$. Fig. 2 illustrates the amount of material that is transported into orbit for the parameter range $0.4 < b < 0.8$ (see figure B2 in Canup (2008) for more details). It shows that a small impact parameter b reduces the material ejected into orbit significantly. The same holds for a reduction of γ, because those collisions are more grazing. Based of simple arguments, we use the limits above to identify the moon forming collision in the (v, b)-plane. Details on how this assumption affects the result are given in Section 4.

Ida et al. (1997) and Kokubo et al. (2000) studied the formation of a moon in a circumplanetary disk through N-body simulations. They found that the final satellite mass scales linearly with the specific angular momentum of the disk. The fraction of the disk material that is finally incorporated into the satellite ranges form 10–55%. Thus, we assume that not more than half of the disk material is accumulated into a single satellite. The angular momentum of the disk is unknown and we can not use the more exact relationships.

3.2. Spin–orbit resonance

Spin–orbit resonance occurs when the spin precession frequency of a planet is close to one of the planet's orbital precession frequencies. It causes large variation in the obliquity (Laskar, 1996), the angle between this spin axis and the normal of the planet's orbital plane. An obliquity stabilizing satellite increases the spin precession frequency to a non-resonant (spin–orbit) regime. To ensure this, one can set a rough limit on the system parameters (Atobe et al., 2004) through

$$\frac{m_s}{a_s^3} \gg \frac{m_*}{a_p^3}, \tag{6}$$

where m_s is the mass of the satellite, m_* the mass of the central star and a_s the semi-major axis of the satellite's orbit and a_p the semi-major axis of the planet.

If the left term of the inequality is much larger than the right one, the spin precession frequency of the planet should be high enough to ensure that it is over the upper limit of the orbital precession frequency so that spin–orbit resonance does not occur. Although this inequality is very simplified, we try to estimate the minimum mass of a satellite such that it is able to stabilize the obliquity of its planet.

The exact semi-major axis of a satellite after formation is unknown but the Roche limit is the lower bound of its semi major axis. The Roche limit a_{R} of the planet is (Murray and Dermott, 1999):

$$a_{\text{R}} = r_s \left(\frac{3m_p}{m_s}\right)^{\frac{1}{3}}, \tag{7}$$

where r_s is the radius of the satellite and m_s its mass and the mass of the planet is given by m_p. Ohtsuki (1993) and Canup and Esposito (1996) provided detailed analytic treatments of the accretion process of satellites in an impact-generated disk. Based on those studies, Kokubo et al. (2000) have shown that the true value of the radius of satellite accretion will not diverge much from the Roche radius in the case of the Earth–Moon system, a typical satellite orbit semi-major axis in their simulations was $a \simeq 1.3 a_{\text{R}}$. We used this approximation to estimate a lower bound on the satellite–planet mass ratio. We rewrite the Roche limit as

$$a_{\text{R}} = \left(\frac{3}{2}\right)^{\frac{2}{3}} \left(\frac{m_p}{\pi \rho_p}\right)^{\frac{1}{3}}, \tag{8}$$

where we used $r_s^3 = \left(\frac{4\pi}{3}\rho_s\right)^{-1} m_s$ and the fact that $\rho_p = \rho_s$ in our model. We insert this in Eq. (6) instead of a_s and get the condition

$$\frac{m_s}{m_p} \gg \frac{9m_*}{4\pi\rho_p a_p^3}. \tag{9}$$

Inserting a density of 2 g cm^{-2}, the density of the bodies in the Morishima simulations, and planet semi-major axis of 1 AU, we get a mass ratio of $\sim 10^{-5}$, which is smaller than the minimum ratio of the smallest and largest particles in our simulations. This is a lower limit on the stabilizing satellite mass but the tidal evolution of the planet–satellite system can alter this limit dramatically.

The orbital precession frequencies of a planet depend on the neighboring or massive planets in its system. In the case of the Earth, Venus, Jupiter and Saturn cause the most important effects. To keep the spin precession frequency high enough, a spin period below 12 h would have the same effect as the present day Moon (Laskar and Robutel, 1993). Hence, even without a massive satellite, obliquity stabilization is possible as long as the planet is spinning fast enough. However, the moon forming impact provides often a significant amount of angular momentum. A more sophisticated analysis including the precession frequencies of the all planets involved or formed in the simulations and a better treatment of the collisions to provide better estimates of the planetary spins would clearly be an improvement but this is out of the scope of this work.

3.3. Tidal evolution

After its formation, even a small satellite is stabilizing the planet's obliquity. Its fate is mainly controlled by the spin of the planet, the orientation of the spin axis of the planet relative to its orbit around the central star and relative to the orbital plane of the satellite and by possible spin–orbit resonance. Which satellites will continue to stabilize the obliquity as they recede from the planet? Orbital evolution is a complicated issue (Atobe and Ida, 2006) and there is still an ongoing debate even in the case of the Earth–Moon system, as mentioned. To classify the different orbital evolutions, it is helpful to introduce the synchronous radius, at which a circular orbital period equals the rotation period of the planet. In the

prograde case, a satellite outwards of the synchronous radius will recede from the planet as angular momentum is tidally transferred from the planet to the satellite and the spin frequency of the planet decreases. In this case the synchronous radius will grow till it equals the satellite orbit. Angular momentum is transferred faster if the mass of the satellite is large, since the tidal response in the planet due to the satellite is greater. Large mass satellites will quickly reach this final co-rotation radius, where their recession stops. Even though their orbital radius becomes larger, these moons are massive enough to satisfy inequality (6) and avoid spin–orbit resonances. Small satellites will recede very slowly compared to their heavy brothers and eventually fulfill the condition (6) within the host star's main sequence life time. In contrast, low mass satellites that form in situ far outside the Roche limit, should this be possible, will probably not stabilize the spin axis. Intermediate mass satellites with $m_s \sim m_{\rm Moon}$ may recede fast enough so that they tend to lose their obliquity stabilizing effect during a main sequence life time. The Earth–Moon system shows that even in this intermediate mass regime, long term stability can occur, since the Moon has stabilized the Earth's obliquity for billion of years.

On the other hand, a satellite inside the synchronous orbit will start to spiral towards the planet, while its angular momentum is transferred to the planet. Soon, it is disrupted by tidal forces or will crash on the planet. How do we know if a satellite forms inside or outside the synchronous radius? The Roche limit (7) depends on the mass and size of the bodies while the synchronous radius $r_{\rm sync}$ is a function of the planet mass m_p and depends inversely proportional on its rotation frequency, which is obtained from equating the gravitational acceleration and the centripetal acceleration:

$$r_{\rm sync} = (Gm_p)^{\frac{1}{3}}\omega_p^{-\frac{2}{3}}. \tag{10}$$

Equating this formula with (7) gives:

$$\left(\frac{3}{2}\right)^{\frac{2}{3}}\left(\frac{m_p}{\pi\rho_p}\right)^{\frac{1}{3}} = (Gm_p)^{\frac{1}{3}}\omega_p^{-\frac{2}{3}}. \tag{11}$$

The planet mass drops out and we get an lower limit on the planet angular velocity to guarantee a satellite outside the Roche radius:

$$\omega_{p,min} = \frac{2}{3}(\pi G\rho_p)^{\frac{1}{2}} = 0.00043 s^{-1}, \tag{12}$$

which equals a rotation period of 4 h. The final planets of Morishima et al. (2010) have generally a high rotational speed, some of them rotating above break up speed. The rotation period after the moon forming collision is usually around 2–5 h, see Fig. 3. A more exact particle growth model without the assumption of perfect sticking would lower the rotation frequency by roughly 30% (Kokubo and Genda, 2010), where we can also include the loss of rotational angular momentum due to satellite formation (full circles). On the other hand, the mean distance of satellite formation is around $1.3a_{\rm R}$, and the maximum rotation period of the planet for a receding satellite changes to ~6 h (dashed line). Hence, applying this to our final sample, 1/4 of all moon forming collisions are excluded.

Finally, for the remaining events, we assume that the planet spin is large enough so that the synchronous radius is initially smaller than the Roche radius. Satellites that form behind the Roche limit will start to recede from the planet. We can conclude that almost every satellite–planet system in our simulation will fulfill (6).

A special outcome of the tidal evolution of a prograde planet–satellite system is described by Atobe and Ida (2006). If the initial obliquity θ of the planet after the moon forming impact is large, meaning that the angle between planet spin axis and planet orbit normal is close to 90°, results in a very rapid evolution when

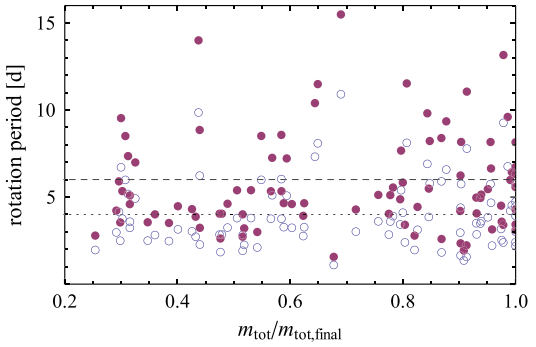

Fig. 3. The mass of the planet in units of its final mass $m_{\rm tot}/m_{\rm tot,final}$ relative to the rotational period in days of the body after the moon-forming collision. This is the final sample but without excluding moon formation inside the synchronous radius. *Empty circles:* The decrease of the planet spin due to escaping material or the formation of a satellite are not included in the calculation of the rotation period, it is based on the perfect accretion assumption. *Full circles:* The angular momentum is assumed to be 30% smaller due to a more realistic collision model (Kokubo and Genda, 2010). The dotted line gives the position of the threshold for synchronous rotation for the Roche radius $a_{\rm R}$, the dashed line gives its position for the $1.3a_{\rm R}$. We want to focus on the better estimate of the angular momentum (full circles) and the initial moon radius (dashed line). Hence, roughly a quarter of all moon forming collisions are excluded in the final sample, Fig. 7.

compared to the previously discussed case of a moon receding to the co-rotating radius and becoming tidally locked. The spin vectors of the protoplanets are isotropically distributed after the giant impact phase (Agnor et al., 1999) and the obliquity distribution corresponds to $p(\theta) = \frac{1}{2}\sin(\theta)$. In the most extreme scenario, a massive prograde satellite will crash onto the planet in a timescale of order 10,000 years after formation, even though it initially recedes. Hence, without a favorable initial obliquity is an added requirement for the survivability of a close or massive moon. We include this by discarding massive and highly oblique impacts by an approximated inequality, based on area B in Fig. 13 in Atobe and Ida (2006):

$$\theta < \frac{\pi}{2} - \frac{\pi}{0.2}\frac{m_s}{m_p}. \tag{13}$$

More exactly, this holds best for planets with $a_p \sim 1$ AU, and it has only to be taken into account if $m_s/m_p < 0.05$. The distribution of the a_p of the simulated planets ranges for 0.1–4 AU. In the case of a smaller distance to the host star, angular momentum is removed faster from the planet–satellite system and the evolution timescales are shorter in general. To use this limit properly, a more general expression for different semi-major axes has to be derived, which is out of the scope of this work. Moreover, it depends linearly on the satellite mass which is overestimated in general. A smaller satellite mass would reduce the number of excluded events in general. However, when applying this constraint on the data, only one out of seven moon forming collisions are affected. Since it is over-simplified, we exclude it from our analysis.

For retrograde impacts, where the impactor hits in opposition to the target's spin, two cases result in differing evolution. If the angular momentum of the collision is much larger than the initial rotational angular momentum, the spin direction of the planet is reversed and any impact-generated disk will rotate in the same direction as the planet's spin. On the other hand, a retrograde collision with small angular momentum will not alter the spin direction of the planet significantly and it becomes possible to be left with a retrograde circumplanetary disk (Canup, 2008). After accretion of a satellite, tidal deceleration due to the retrograde protoplanet spin will reduce the orbital radius of the bodies continuously till they merge with the planet. Hence, a long-lived satellite can hardly form in this case.

In order to find a reasonable threshold between these two regimes based on the limited information we have, we assume that the sum of the initial angular momenta of the bodies $\vec{L}_t$ and $\vec{L}_i$ and of the collision $\vec{L}_{col}$ equals the spin angular momentum of the planet and the satellite $\vec{L}_{planet}$ and $\vec{L}_{moon}$ after the collision plus the angular momentum of the orbiting satellite $\vec{L}_{orbit}$ at $1.3a_R$ parallel to the collision angular momentum.

Comparing initial and final angular momenta gives:

$$\vec{L}_{col} + \vec{L}_t + \vec{L}_i = \vec{L}_{orbit} + \vec{L}_{moon} + \vec{L}_{planet}, \tag{14}$$

where $\vec{L}_{orbit}$ is parallel to $\vec{L}_{col}$:

$$\vec{L}_{orbit} = |\vec{L}_{orbit}| \frac{\vec{L}_{col}}{|\vec{L}_{col}|}. \tag{15}$$

We assume that $|\vec{L}_{moon}|/|(\vec{L}_{orbit} + \vec{L}_{planet})| \ll 1$ and

$$L_{orbit} = m_s a_s^2 n = m_s a_s^2 \sqrt{\frac{Gm_p}{a_s^3}} = m_s \sqrt{G a_s m_p}, \tag{16}$$

where n is the orbital mean motion of the satellite if its mass is much smaller than the planet mass, $n \sim \sqrt{Gm_p/a_s^3}$, and G is the gravitational constant. With $a_s = a_R$ we get

$$L_{orbit} = m_s (Gm_p)^{\frac{1}{2}} \left(\frac{3}{2\pi} \frac{1}{\rho_p} m_p \right)^{\frac{1}{6}}. \tag{17}$$

If

$$\frac{|\vec{L}_{col} + \vec{L}_t|}{\sqrt{(L_{col}^2 + L_t^2)}} < 1, \tag{18}$$

the collision is retrograde. Hence,

$$\vec{L}_{planet} = \vec{L}_{col} + \vec{L}_t + L_i - |\vec{L}_{orbit}| \frac{\vec{L}_{col}}{|\vec{L}_{col}|}, \tag{19}$$

If $\vec{L}_{planet}$ and $\vec{L}_{orbit}$, which is parallel to $\vec{L}_{col}$, are retrograde,

$$\frac{|\vec{L}_{col} + \vec{L}_{orbit}|}{\sqrt{(L_{col}^2 + L_{orbit}^2)}} < 1, \tag{20}$$

the satellite will be tidally decelerated. Those cases are excluded from being moon-forming events. Spin angular momenta are in general overestimated since material is lost during collisions in general, but as mentioned above, the simulations of Morishima et al. (2010) assume perfect accretion. We include this consideration by reducing the involved spins and the orbit angular momentum of the satellite by 30%, (Kokubo and Genda, 2010).

Both scenarios described above, a large initial obliquity or a retrograde orbiting planet, might not become important until subsequent impactors hit the target. Giant impacts can change the spin state of the planet in such a way that the satellite's fate is to crash on the planet. This scenario is discussed in the next section.

3.4. Collisional history

We exclude all collisions from being satellite-forming impacts whose target is not one of the final planets of a simulation. A satellite orbiting an impact is lost through the collision with the larger target.

Multiple giant impacts occur during the formation process of a planet and it is useful to study the impact history in more detail. Subsequent collisions and accretion events on the planet after the satellite-forming event may have a large effect on the final outcome of the system. We choose a limit of $5m_{planetesimal}$ to distinguish between large impacts and impacts of small particles, which are responsible for the ordered accretion. To stay consistent, the same limit is used below to exclude small impactors from our analysis. We divide all identified moon forming events into four groups:

(a) The moon forming event is the last major impact on the planet. Subsequent mass growth happens basically through planetesimal accretion (see merger tree at the bottom in Fig. 1).
(b) There are several moon forming impacts, in which the last impact is the last major impact on the planet.
(c) The moon forming event is not the last giant impact on the planet. The satellite can be lost due to a disruptive near or head-on-collision (Stewart and Leinhardt, 2009) of an impactor and the satellite. A late giant impact on the planet can change the spin axis of the planet and the existing satellite can get lost due to tidal effects (see tree at the top in Fig. 1).
(d) There are several moon forming events in the impact history of the planet, followed by additional major impacts. As before, the moon forming collisions can remove previously formed satellites. On the other hand, and existing moon

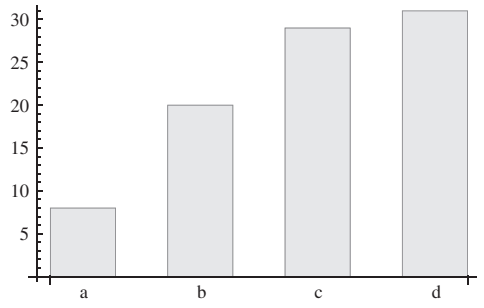

Fig. 4. This bar chart shows the distribution of the moon forming collisions in four groups with different impact history with respect to the last major impact. (a) The moon forming impact is the last major impact on the planet. (b) There are multiple moon forming impacts, but the last one is also the last major impact. (c) There is only one moon forming impact and it is followed by subsequent major impacts. (d) There are multiple moon forming impacts, but the last of them is followed by subsequent major impacts.

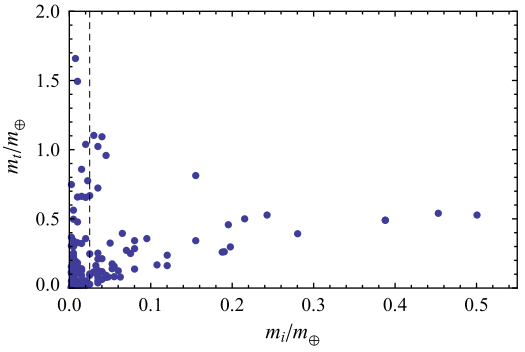

Fig. 5. The target mass m_t versus the impactor mass m_i involved in the satellite-forming collisions. We exclude events that include impactors and targets with masses of the order of the initial planetesimal mass (m_i, $m_t < 5m_{planetesimal}$) indicated by the dashed line to avoid resolution effects. The line in the case of the target mass is not shown since it is very close the frame. Due to this cut, the total number of satellites decreases significantly while the number of massive satellites in our analysis increases. The shown sample is the final set of events without applying the threshold for small particles. Therefore, the number of accepted events in this plot does not equal the number of satellites in Fig. 7, since this cut is not applied. If the cut is used, new events are accepted, which where neglected before because they were not the last moon forming impacts on the target.

can have an influence on the circumplanetary disk formed by a giant impact and can suppress the formation of multiple satellites orbiting the planet.

The final states of the planet–satellite systems in group *c* are difficult to estimate, since such systems might change significantly by additional giant impacts. To a lesser extent, this holds for *d*, but those systems are probably more resilient to the loss of satellites by direct collisions. The number of events per group is shown in Fig. 4. Group *c* and *d* include more then 2/3 of all collisions. Group *c* is the most uncertain and we use it to quantify the error on the final sample.

Furthermore, we exclude impactors and targets that have masses of the order of an initial planetesimal mass ($5m_{\mathrm{planetesimal}} \sim 0.0025m_{\oplus}$) from producing satellite forming events since their masses are discretized and related to the resolution of the simulation. In addition, small impactors will probably not have enough energy to eject a significant amount of material into a stable orbit. Therefore, setting a lower limit on the target and impactor mass results in the exclusion of many collisions but not in a significant underestimation of the true number of satellites, see Fig. 5.

4. Uncertainties

Our final result depends on several assumptions, limitations and approximations. In this section we want to quantify them as much as possible and summarize them.

The data of Morishima et al. (2010) we are using has two peculiarities worth mentioning: The focus on the Solar System and the small number of simulations per set of initial conditions.

The simulations were made in order to reproduce the terrestrial planets of the Solar System. The central star has 1 solar mass and the two gas giants that are introduced after the gas dissipation time scale have the mass of Jupiter and Saturn and the same or similar orbital elements. Although the initial conditions like initial disk mass or gas dissipation time scale are varied in a certain range, the simulated systems do not represent general systems with terrestrial planets. However, we assume that the range of impact histories is representative of the range that would be seen in other systems. Merger trees (Fig. 1) reveal that those simulations cover a huge diversity of impact histories. But future work will need to investigate the full range of impact histories that could be relevant to the formation of terrestrial planets in other, possibly more exotic, extra-Solar Systems.

However, the set of simulation covers a broad range of initial conditions. But for every set of initial conditions, only one simulation exists since they are very time consuming (each simulation requires about 4 months of a quad-core CPU). Therefore, it is difficult to separate effects of the choice of certain initial parameters from effects of stochastic processes. We grouped our moon forming events with respect to the simulation parameter in question. The only parameter that reveals an effect on the final sample is the gas dissipation time scale τ_{gas}. Larger time scales lead to less moon forming collisions. This variation is correlated with mass and number of final planets. If the gas disk stays for several million years, the bodies are affected by the gas drag for a long time, spiral towards the Sun and get destroyed. Therefore, there are less giant impacts and smaller planets. One would suppose that the initial mass of the gas disk should be correlated to the mass of the final planets and therefore to the number of giant impacts, but the two initial protoplanetary disk masses of $5m_{\oplus}$ and $10m_{\oplus}$ show no essential difference.

The approach we use to identify events and estimate the mass of the satellite is also based on various approximations and limitations.

4.1. Satellite mass

The method we use to calculate the mass of the circumplanetary disk is valid to better than a factor of 2 within the parameter range we use (Canup, 2008). Only 10–55% of the disk mass are embodied in the satellite (Kokubo et al., 2000). Therefore, the mass we use is just half of the estimated disk mass, and in the worst case, the satellite is ten times less massive than estimated. This uncertainty affects the number of massive satellites but not the number of satellites in general.

In the (v, b)-plane, we use the same restrictive limits to constrain the collision events. Fig. 6 presents all moon forming events and this parameter range. Hence, this area gives just a lower bound. We see that it covers some of the most populated parts, but a significant amount of the collisions are situated outside this area. Eq. (1) can help to constrain the outcome of the collisions close outside of the shaded region. A collision with a small impact parameter ($b < 0.4$) will bring very little material into orbit, whatever the impactor mass involved. In this regime, we have few events with high impact velocity and even with high velocity it might be very hard to eject a significant amount of material into orbit. Hence, we exclude them from being moon forming events. For intermediate impact parameter ($0.4 < b < 0.8$), there are high velocity events ($v > 1.4$). Above a certain velocity threshold, depending on b and γ, most material ejected by the impact will escape the system and the disk mass might be too small to form a satellite of interest. A large parameter ($b > 0.8$) describes a highly grazing collision. It is difficult to extent Eq. (1) for larger b, since these collisions will probably result in a hit-and-run events for high velocities. SPH simulations (Canup, 2004, 2008) suggest that high relative impact velocities ($v > 1.4$) will increase rapidly the amount of material that escapes the system. Nevertheless, these sets of simulations focus not on general impacts and the multi-dimensional collision parameter space is not studied well enough to describe those collisions in more detail. Detailed studies of particle collisions will hopefully be published in the near future (e.g. Kokubo and Genda, 2010). Based on the arguments above, the

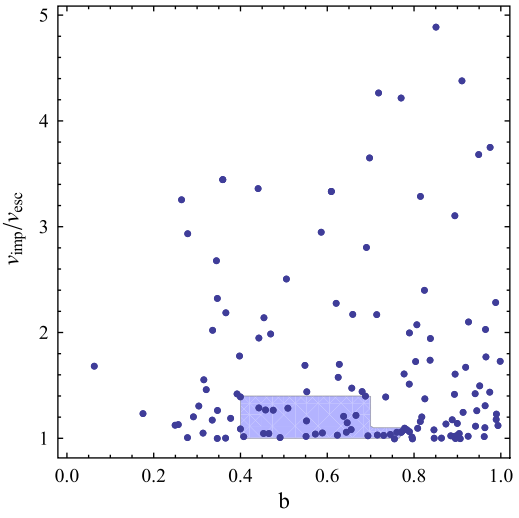

Fig. 6. The parameter space in the (b, $v_{\mathrm{imp}}/v_{\mathrm{esc}}$)-plane. The darker region gives the region for with Eq. (1) holds up to a factor of 2. The shown sample is the final sample without the constraints in the (b, v)-plane. Similar to Fig. 5, the number of accepted events will increase when applying the cut, since an earlier moon forming impact on the planet can become suitable. As one expects, there are more collisions with larger impact parameter. Ignoring events with smaller b or higher v (see text), there is still a significant group of events with $b > 0.8$. Much material escapes in the high velocity cases and no moon will form, but the low velocities are hard to exclude or calculate.

events inside of the shaded area form our final sample. Most of the collisions outside the area will not form a moon. The collisions with $0.4 < b < 0.8$ and velocity slightly above $v = 1.4$ and with $b > 0.8$ and small velocity ($v < 1.4$) are events with unknown outcome. Including those events, the final number of possible moon forming events is increased by not more than a factor of 1.5.

4.2. Tidal evolution – rotation period

In order to separate receding satellites from satellites which are decelerated after their formation, we check if the initial semi-major axis of the satellite is situated outside or inside of the synchronous radius of the planet. In Fig. 3, the sample is plotted twice, once including a general correction for angular momentum loss due to realistic collisions and once without correction. Moreover, two thresholds for the synchronous radius are shown. We choose the threshold at $1.3a_R$ (dashed line) and the corrected rotation period (full circles) to be the most justified case. There are two most extreme cases: the threshold situated at $1a_R$ (dotted line) and a corrected rotation period (full circles) indicates that only one in five collisions lead to receding moons. On the other hand, the threshold situated at $1.3a_R$ and a rotation period directly obtained form the simulations (empty circles) indicates that only one in eight collisions lead to a non-receding moon.

4.2.1. Tidal evolution – retrograde satellites

To exclude retrograde orbiting satellites, we use a simple relation between the angular momenta involved in a collision. Since we have no exact data of the angular momentum distribution after the collision, this limit is very approximate. To get an estimate of the quality of this angular momentum argument, we study the effect of this threshold on the final sample. First, roughly half of all moon-forming impacts are retrograde. But in almost every case, L_{orbit} is much smaller than the initial angular momentum. Only four of the retrograde collisions have not enough angular momentum to provide a significant change of the spin axis and those planet–satellite systems remain retrograde. Therefore, our final estimate is not very sensitive to this angular momentum argument. A larger set of particle collision simulations could provide better insight into the angular momentum distribution.

4.2.2. Collision history

Two issues affecting the collision history can change the final result. These being the mass cut we choose to avoid resolution effects, where we exclude small impactors and targets, and the uncertainty about the final state of the planet–satellite system because of multiple and subsequent major impacts. The first limit seems to be well justified. If we change the threshold mass or exclude the cut completely, we change mainly the number of very small, non-satellite forming, impacts since in these cases there is usually insufficient energy to bring a significant amount of material into orbit. The second issue has a significant effect on the final sample, however. Group *c*, the single moon forming collisions followed by large impacts, contains the most uncertain set of events: Neither all satellites survive the subsequent collisions nor is it likely that all satellites get lost. Hence, in the extreme case, almost half of all initially formed moons, all of Group *c*, could be lost and our final result reduced significantly.

An overview of the above uncertainties in given in Table 1. Additional planet formation simulations are necessary to quantify a large part of existing uncertainties. Simulations of protoplanet collisions exploring the multi-dimensional collision parameter space are desirable and will hopefully be published soon and might help to constrain the parameters for moon formation better. A study of the effect of subsequent accretion of giant bodies after a moon forming impact might give better insight in the evolution of a planet–satellite system.

Table 1
A list of the different conditions that affect the final number of satellites. The uncertainty factor gives the range in which the final result varies around the most justified. 1 Equals the final value. The first two factors can not be estimated, it shows that additional simulations would be helpful. The estimation of the satellite mass is separated from the rest of the list, since the uncertainty on this estimate does not change the number of satellites in the final sample, in contrast to the others. The exclusion of retrograde satellites is very approximate, but it has almost no effect on the final sample and therefore this factor is close to 1.

	Uncertainty factor
Range of initial parameters	High
Number of simulations	High
Collision parameter	1–1.5
Tidal evolution – rotation	0.5–2
Tidal evolution – retrograde orbit	~1
Collision history	0.5–1
Satellite mass – disk mass	0.5–2
Satellite mass – accretion efficiency	0.2–1

5. Discussion and results

Under these restrictive conditions we identify 88 moon forming events in 64 simulations, the masses of the resulting planet–satellite systems are shown in Fig. 7. On average, every simulation gives three terrestrial planets with different masses and orbital characteristics and we have roughly 180 planets in total. Hence, almost one in two planets has an obliquity stabilizing satellite in its orbit. If we focus on Earth–Moon like systems, where we have a massive planet with a final mass larger than half of an Earth mass and a satellite larger than half a Lunar mass, we identify 15 moon forming collisions. Therefore, 1 in 12 terrestrial planets is hosting a massive moon. The main source of uncertainties results from the modelling of the collision outcomes and evolution of the planet–satellite system as well as the small number of simulation and the limited range of initial conditions. We do not include the latter in our estimate. Hence, we expect the total number of Earth–Moon like systems in all our simulations to be in a range from 4 to 45. This results in a low-end estimate of 1 in 45 and a high-end estimate of 1 in 4. In addition, taking into account the uncertainties on the estimation of the satellite mass, roughly 60 of those systems are formed in the best case or almost no such massive satellites are formed if the efficiency of the satellite accretion in the circumplanetary disk is very low.

There are several papers, where the authors performed N-body simulations and searched for moon-forming collisions. Agnor et al.

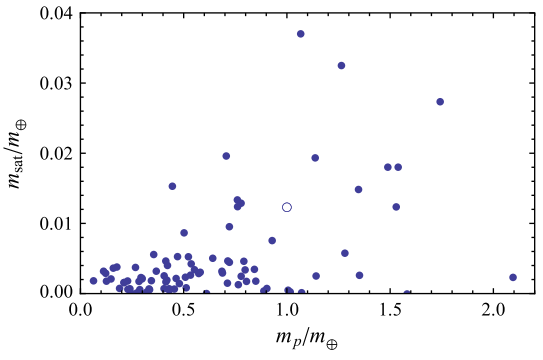

Fig. 7. The masses of the final outcomes of the planets for which we identified satellite forming collisions. $m_{satellite}$ is the mass of the satellite, assuming an accretion efficiency of 50%, and m_p is the mass of the planet after the complete accretion. The circle indicates the position of the Earth–Moon system with the assumption $m_{disk} = m_{Moon}$.

(1999) started with 22–50 planetary embryos in a narrow disk centered at 1 AU. They estimated around 2 potentially moon-forming collisions per simulation, where the total angular momentum of the encounter exceeds the angular momentum of the Earth–Moon system. They pointed out that this number is somewhat sensitive to the number, spacing, and masses of the initial embryos. O'Brien et al. (2006) performed simulations with 25 roughly Mars-mass embryos embedded in a disk of 1000 non-interacting (with each other) planetesimals in an annulus from 0.3 to 4.0 AU. They found that giant impact events which could form the Moon occur frequently in the simulations. These collisions include a roughly Earth-size target whose last large impactor has a mass of 0.11–$0.14M_E$ and a velocity, when taken at infinity, of 4 km/s, as found by Canup (2004). O'Brien et al. pointed out that their initial embryo mass is close to the impactor mass. Raymond et al. (2009) set up about 90 embryos with masses from 0.005 to $0.1M_E$ in a disk of more than 1000 planetesimals, again with non-interaction of the latter. Assuming again Canup's requirements ($v/v_{\rm esc} < 1.1$, $0.67 < \sin\theta < 0.76$, $0.11 < \gamma < 0.15$), only 4% of their late giant impacts fulfill the angle and velocity criteria. They concluded that Earth's Moon must be a cosmic rarity but a much larger range of late giant collisions would produce satellites with different properties than the Moon. The initial embryo size seems to play a role in those results. In contrast, the simulations of Morishima et al. (2010) start with 2000 fully interacting planetesimals and since embryos form self-consistently out of the planetesimals in these simulations, problems with how to seed embryos are completely avoided. To constrain the simulations, Morishima et al. (2010) were interested in the timing of the Moon-forming impact. To identify potentially events, they searched for a total mass of the impactor and the target $>0.5m_{\oplus}$, a impactor mass $>0.05m_{\oplus}$ and a impact angular momentum $>L_{\rm EM}$. They found almost 100 suitable impacts in the 64 simulations. Since their sample also includes high velocity or grazing impacts, although constrained through the more general angular momentum limit, and does not take into account collision history and tidal evolution, the difference to our result is not surprising.

Life on planets without a massive stabilizing moon would face sudden and drastic changes in climate, posing a survival challenge that has not existed for life on Earth. Our simulations show that Earth-like planets are common in the habitable zone, but planets with massive, obliquity stabilizing moons do occur only in 10% of these.

Acknowledgments

We thank David O'Brien and an anonymous reviewer for many helpful comments. We thank University of Zurich for the financial support. We thank Doug Potter for supporting the computations made on zBox at University of Zurich.

References

Agnor, C.B., Canup, R.M., Levison, H.F., 1999. On the character and consequences of large impacts in the late stage of terrestrial planet formation. Icarus 142, 219–237.
Atobe, K., Ida, S., 2006. Obliquity evolution of extrasolar terrestrial planets. Icarus 188, 1–17.
Atobe, K., Ida, S., Ito, T., 2004. Obliquity variations of terrestrial planets in habitable zones. Icarus 168, 223–236.
Berger, A. (Ed.), 1989. Climate and Geo-Science – A Challenge for Science and Society in the 21st Century. Kluwer Academic, Dordrecht.
Berger, A.L., Imbrie, J., Hays, J., Kukla, G., Saltzman, B. (Eds.), 1984. Milankovitch and Climate – Understanding the Response to Astronomical Forcing. D. Reidel, Norwell.
Cameron, A.G.W., Benz, W., 1991. The origin of the Moon and the single-impact hypothesis IV. Icarus 92, 204–216.
Cameron, A.G.W., Ward, W.R., 1976. The origin of the Moon. Proc. Lunar Sci. Conf. 7, 120–122.
Canup, M.R., 2004. Simulations of a late lunar-forming impact. Icarus 168, 433–456.
Canup, M.R., 2008. Lunar-forming collisions with pre-impact rotation. Icarus 196, 518–538.
Canup, M.R., Esposito, L.W., 1996. Accretion of the Moon from an impact-generated disk. Icarus 119, 427–446.
Chamberlin, T.C., 1905. In Carnegie Institution Year Book 3 for 1904. Carnegie Inst., Washington, DC, pp. 195–234.
Chambers, J.E., 1999. A hybrid symplectic integrator that permits close encounters between massive bodies. Mon. Not. R. Astron. Soc. 304, 793–799.
Chambers, J.E., Wetherill, G.W., 1998. Making the terrestrial planets: N-body integrations of planetary embryos in three dimensions. Icarus 136, 304–327.
Dones, L., Tremaine, S., 1993. On the origin of planetary spin. Icarus 103, 67–92.
Duncan, M., Levison, H.F., Lee, M.H., 1998. A multiple time step symplectic algorithm for integrating close encounters. Astron. J. 116, 2067–2077.
Goldreich, P., Peale, S.J., 1970. The obliquity of Venus. Astron. J. 75, 273–284.
Halliday, A.N., 2000. Terrestrial accretion rates and the origin of the Moon. Earth Planet. Sci. Lett. 176, 17–30.
Hartmann, W.K., Davis, D.R., 1975. Satellite-sized planetesimals and lunar origin. Icarus 24, 504–515.
Ida, S., Canup, R.M., Stewart, G.R., 1997. Lunar accretion from an impact-generated disk. Nature 389, 353–357.
Kokubo, E., Genda, H., 2010. Formation of terrestrial planets from protoplanets under a realistic accretion condition. Astrophys. J. 714, 21–25.
Kokubo, E., Ida, S., 1998. Oligarchic growth of protoplanets. Icarus 131, 171–178.
Kokubo, E., Ida, S., Makino, J., 2000. Evolution of a circumterrestrial disk and formation of a single Moon. Icarus 148, 419–436.
Kokubo, E., Kominami, J., Ida, S., 2006. Formation of terrestrial planets from protoplanets. I. Statistics of basic dynamical properties. Astrophys. J. 642, 1131–1139.
Laskar, J., 1996. Large scale chaos and marginal stability in the Solar System. Celest. Mech. Dynam. Astron. 6, 115–162.
Laskar, J., Robutel, P., 1993. The chaotic obliquity of the planets. Nature 361, 608–612.
Laskar, J., Joutel, F., Robutel, P., 1993. Stabilization of the Earth's obliquity by the Moon. Nature 361, 615–617.
Lathe, R., 2004. Fast tidal cycling and the origin of life. Icarus 168, 18–22.
Lathe, R., 2006. Early tides: Response to Varga et al.. Icarus 180, 277–280.
Lissauer, J.J., 1993. Planet formation. Ann. Rev. Astron. Astrophys. 31, 129–174.
Milankovitch, M., 1941. Kanon der Erdbestrahlung und seine Anwendung auf das Eiszeitproblem. Kanon Koeniglich Serbische Academie Publication, p. 133.
Morishima, R., Stadel, J., Moore, B., 2010. From planetesimals to terrestrial planets: N-body simulations including the effects of nebular gas and giant planets. Icarus 207, 517–535.
Murray, C.D., Dermott, S.F., 1999. Solar System Dynamics. Cambridge University Press, Cambridge.
O'Brien, D.P., Morbidelli, A., Levison, H.F., 2006. Terrestrial planet formation with strong dynamical friction. Icarus 184, 39–58.
Ohtsuki, K., 1993. Capture probability of colliding planetesimals: Dynamical constraints on accretion of planets, satellites, and ring particles. Icarus 106, 228–246.
Raymond, S.N., Quinn, Th., Lunine, J.I., 2004. Making other earths: Dynamical simulations of terrestrial planet formation and water delivery. Icarus 168, 1–17.
Raymond, S.N., O'Brien, D.P., Morbidelli, A., Kaib, N.A., 2009. Building the terrestrial planets: Constrained accretion in the inner Solar System. Icarus 203, 644–662.
Safronov, V.S., 1969. Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets. Moskau, Nauka. Engl. Transl. NASA TTF-677, 1972.
Stadel, J., 2001. Cosmological N-body Simulations and their Analysis. Ph.D. Thesis, University of Washington.
Stewart, S.T., Leinhardt, Z.M., 2009. Velocity-dependent catastrophic disruption criteria for planetesimals. Astrophys. J. 691, 133–137.
Touboul, M., Kleine, T., Bourdon, B., Palme, H., Wieler, R., 2007. Late formation and prolonged differentiation of the Moon inferred from W isotopes in lunar metals. Nature 450, 1206–1209.
Varga, P., Rybicki, K.R., Denis, C., 2006. Comment on the paper Fast tidal cycling and the origin of life by Richard Lathe. Icarus 180, 274–276.
Ward, W.R., 1974. Climate variations on Mars. 1. Astronomical theory of insolation. J. Geophys. Res. 79, 3375–3386.
Ward, P.D., Brownlee, D., 2000. Rare Earth: Why Complex Live is Uncommon in the Universe. Copernicus. Springer-Verlag, New York.
Ward, W.R., Rudy, D.J., 1991. Resonant obliquity of Mars? Icarus 94, 160–164.
Wetherill, G.W., Stewart, G.R., 1989. Accumulation of a swarm of small planetesimals. Icarus 77, 330–357.

4

Paper II
Elemental Abundances

In this chapter, we explored the elemental composition of planets. This work was motivated by the fact that the bulk composition of exoplanets has not been studied extensively before, although elemental abundances in planets provide the initial conditions for life and contain much information about the origin of the building blocks of the planets. The pioneering work of Bond et al. (2010) [116] was our starting point. It presents a method that traces how solids that condense in the protoplanetary disk are comprised in the final planets. Since this method divides planet formation in two stages, the formation of solids according to the thermodynamics in the disk and the formation of planets via the coagulation of planetesimals, there are different ways to treat the transition from one stage to the other as well as many model parameters. It is therefore reasonable to study the attempt of Bond et al. (2010) concerning the dependence on the different assumptions.

We observed that the estimated elemental composition of planets depends significantly on the model that describes the thermodynamics of the protoplanetary disk. Nevertheless, given that the method spans the whole process of planet formation, the results concerning the Solar system planets are very promising.

This paper was published 2012 in Icarus [117].

Icarus 221 (2012) 859–874

Contents lists available at SciVerse ScienceDirect

Icarus

journal homepage: www.elsevier.com/locate/icarus

On the origin of elemental abundances in the terrestrial planets

S. Elser [a,*], M.R. Meyer [b], B. Moore [a]

[a] Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
[b] ETH Zürich, Institut für Astronomie, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland

ARTICLE INFO

Article history:
Received 20 April 2012
Revised 14 September 2012
Accepted 14 September 2012
Available online 29 September 2012

Keywords:
Abundances, Interiors
Planetary formation
Terrestrial planets

ABSTRACT

The abundances of elements in the Earth and the terrestrial planets provide the initial conditions for life and clues as to the history and formation of the Solar System. We follow the pioneering work of Bond et al. (Bond, J.C., Lauretta, D.S., O'Brien, D.P. [2010]. Icarus 205, 321–337) and combine circumstellar disk models, chemical equilibrium calculations and dynamical simulations of planet formation to study the bulk composition of rocky planets. We use condensation sequence calculations to estimate the initial abundance of solids in the circumstellar disk with properties determined from time dependent theoretical models. We combine this with dynamical simulations of planetesimal growth that trace the solids during the planet formation process and which include the effects of gravitational and hydrodynamical mixing. We calculate the elemental abundances in the resulting rocky planets and explore how these vary with the choice of disk model and the initial conditions within the solar nebula.

Although certain characteristics of the terrestrial planets in the Solar System could be reproduced, none of our models could reproduce the abundance properties of all the planets. We found that the choice of the initial planetesimal disk mass and of the disk model has a significant effect on composition gradients. Disk models that give higher pressure and temperature result in larger variations in the bulk chemical compositions of the resulting planets due to inhomogeneities in the element abundance profiles and due to the different source regions of the planets in the dynamical simulations. We observed a trend that massive planets and planets with relatively small semi-major axes are more sensitive to these variations than smaller planets at larger radial distance. Only these large variations in the simulated chemical abundances can potentially explain the diverse bulk composition of the Solar System planets, whereas Mercury's bulk composition cannot be reproduced in our approach.

1. Introduction

The process of planet formation is far from understood. The established scenario for the formation of the Earth and other terrestrial planets is that most of their masses were built up through the gravitational collisions and interactions of smaller bodies (Chamberlin, 1905; Safronov, 1969; Lissauer, 1993). These smaller bodies, often called planetesimals, form via the coagulation of dust in the midplane of a circumstellar disk. This transition from dust particles to gravitationally interacting bodies is an important topic of ongoing research (Weidenschilling, 1980; Johansen et al., 2007; Chiang and Youdin, 2010).

Many direct simulations of the formation of rocky planets from planetesimals and protoplanets have been performed (Chambers and Wetherill, 1998; Agnor et al., 1999; Raymond et al., 2004; O'Brien et al., 2006; Kokubo et al., 2006; Raymond et al., 2009). One of the goals of these numerical studies is to reproduce the inner part of the Solar System, primarily the number of planets and their mass distribution and orbital parameters. Additional constraints can be the delivery of water rich material onto the planets (Raymond et al., 2004) or the timing of a Moon forming impact taking place (Touboul et al., 2007). Dynamical simulations by Morishima et al. (2010) include gas drag and type-I migration and the secular perturbative effects of Jupiter and Saturn. Some of the simulated planetary systems partially satisfy the main constraints given by our Solar System but there are still some issues to resolve. For example, the small mass of Mars is not reproduced in most of the dynamical simulations, although a possible solution to this was recently proposed by Walsh et al. (2011).

Determining the composition of the terrestrial planets is a challenging task. In case of the Earth, samples from the crust can be easily obtained and lava provides samples from the mantle. There are samples from Mars in the form of meteorites and in situ analyses of rocks from spacecrafts and probes on Venus and Mars. Remote sensing and geophysical measurements provide additional data for all planets. Modeling has to include many different processes that are involved in creating the final bulk chemical composition of a terrestrial planet. It reveals that the bulk

* Corresponding author.
E-mail addresses: selser@physik.uzh.ch (S. Elser), mmeyer@phys.ethz.ch (M.R. Meyer), moore@physik.uzh.ch (B. Moore).

0019-1035/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.icarus.2012.09.016

composition of the planets shows some significant differences in the relative abundance of the major rock forming elements (Morgan and Anders, 1980; Kargel and Lewis, 1993; Lodders and Fegley, 1997) or their metallic core masses, e.g. Righter et al. (2006). Most meteorites and the rocky planets show a depletion in volatile elements relative to primitive carbonaceous CI meteorites, which resulted from the formation of solids in the protoplanetary disk (Davis, 2006) or from volatile loss during the planet formation process. Albarède (2009) pointed out that the volatile depletion of the Earth can be explained by a very hot protoplanetary disk in the inner part of the Solar System, which suppresses the condensation of volatile elements. On the other hand, the Earth is enriched in refractory elements (Palme et al., 2003; Rubie et al., 2011).

When considering the final elemental compositions of the planets it is also important to consider the effects of dynamical evolution. The planetesimals that eventually come together to form a terrestrial planet, themselves formed in various locations in the protoplanetary disk. Their bulk composition can depend strongly on their initial location since the formation of solid species in the protoplanetary disk will most likely depend on the local gas composition, temperature and pressure. Gravitational scattering among the planetesimals and protoplanets and radial movement due to interaction with the gas disk results in a radial mixing of the building blocks of the planets and in different bulk chemical compositions of the final planets.

Bond et al. (2010a) introduced the bulk chemical composition of the Solar System's planets as an additional constraint on the models of their formation. Based on an α-disk model by Hersant et al. (2001), they determined the spatial equilibrium composition of solid material in the inner protoplanetary disk with the HSC Chemistry software package. These spatial solid material abundances were mapped onto the initial planetesimals in the dynamical simulations of O'Brien et al. (2006). The initial location of each planetesimal that is embodied in the final planets is traced in those simulations. Assuming that the bulk composition of the final proto-planets is the sum of the solids embodied in the initial planetesimals, they estimated the bulk chemical composition of the simulated planets. They found that these are comparable to the bulk chemical composition in the Solar System in the case of most elements, although volatiles are an exception and were often overestimated. Furthermore, Bond et al. (2010b) applied their method to extrasolar giant planet systems. They performed dynamical simulations designed to study the formation of rocky planets around those stars with known gas giants detected via radial velocity and known abundances from high resolution stellar spectra. They estimated the bulk chemical composition of the simulated planets and sometimes extreme variations in composition were found.

In this report, we followed the methodology of Bond et al. (2010a). We combined different disk models and the dynamical simulations performed by Morishima et al. (2010) to estimate the bulk chemical composition of the simulated planets. Bond et al. (2010a) were interested in "making the Earth" and reproducing abundances trends, geochemical ratios and oxidation states of the planets. In addition, they studied the late veneer and the loss of volatiles in collisions. In comparison, we focused on the bulk chemical compositions and on variations that might result from different initial conditions in the dynamical simulations and our assumptions in the modeling of the gas disk. In addition, we did not treat the transition from the disk model to the dynamical simulations as a free parameter. Instead, we chose a more self-consistent transition. With those differences, we hope to improve the methodology. In the following, our *disk models* give a time dependent description of the density, temperature and pressure profile of a protoplanetary disk – the physical conditions for the equilibrium chemistry calculations.

The goal of this work consisted of three parts: First, we wished to infer how the distribution of solids in the radial direction in the protoplanetary disk depended on the choice of the disk model. Bond et al. (2010a) is only based on one model, whereas we explored three simple but widely used disk models and their resulting chemical profiles. This allowed us to estimate how much the bulk composition of planets depends on the conditions in the disk.

Second, we focused on the dynamical simulations and tried to understand the link between the initial locations of the planetesimals embodied in simulated planets and the planet's final position and mass. A detailed understanding of this link gives deeper insights into planet formation from the dynamical point of view. We compared the cumulated initial distribution of groups of planets to infer if this can be another indicator of trends in the composition of planets.

Finally, we studied the dependence of the estimated bulk composition of the simulated planets on the disk model and the initial conditions of the dynamical simulations. We were especially interested in the most extreme cases that can be obtained, since such an extreme choice can probably have a strong effect on the results in bulk composition studies. Hence, we could cover a broader range of initial assumptions than Bond et al. (2010a). We compared our results with Solar System bulk compositions to see if this modeling can reproduce earlier predictions.

There are uncertainties in the modeling of the protoplanetary disk and assumptions are made in the determination of the equilibrium composition. The transition from the disk model to the disk of solids used in the dynamical simulations is performed in a simple way. The dynamical simulations represent different sets of initial conditions and their effect on the formation process is only poorly understood. The biggest challenge in this work was to deal with those issues and to investigate how the final results were affected by them. We also keep in mind that fundamental processes like disequilibrium condensation and core–mantle differentiation are not taken into account although they might play a crucial role in determining the bulk chemical composition of a planet.

This report is structured as follows: In Section 2, we review the disk models that are used. We briefly present the tool to calculate the equilibrium abundance in the disk and the dynamical simulations. In Section 3, the bulk chemical compositions of the planets are shown and the differences that arise from the different disk models and initial conditions. In Section 4, the caveats and implications of our approach are summarized. Finally, the most important conclusions are given in Section 5.

2. Methodology

Our method consists of a two-step approach first treating the condensation of solids in an evolving circumstellar disk and then following the reshuffling of those materials dynamically through the planet formation process. The evolution of the surface density, pressure and temperature profiles in the disk are given by a disk model. Based on these profiles, a model that describes the chemical processes in the disk gives the abundance of solid species in the disk midplane in space and time. The disk can evolve until it satisfies a selected criterion based on the initial surface density in the dynamical simulations. At this transition time, the distribution of solids in the disk midplane based on the disk model gives the bulk composition of the initial planetesimals in the dynamical simulations. The following dynamical interactions of the particles and their mergers link the bulk composition of the initial planetesimals to the bulk chemical composition of the final planets.

2.1. Disk models

Observations reveal that protoplanetary disks are not static structures. The observed lifetimes are several millions of years over

which period mass and angular momentum are transported in the radial direction (Hartmann, 1998). The evolution of the disk's surface density Σ follows from the conservation of angular momentum and mass and can be expressed in the evolution equation, which is given in the case of a thin disk by Lynden-Bell and Pringle (1974):

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{r}\frac{\partial}{\partial r}\left(r^{1/2}\frac{\partial}{\partial r}\left(\Sigma \nu r^{1/2}\right)\right). \quad (1)$$

The angular momentum transport is controlled by the viscosity ν. In the widely-used α-prescription, the viscosity is expressed in the form

$$\nu = \alpha c_s H = \alpha c_s^2 \Omega^{-1}, \quad (2)$$

where c_s is the local sound speed and H is the local scale height of the disk. They are related by $c_s = \Omega H$, with the Keplerian angular motion

$$\Omega = \sqrt{\frac{GM_\odot}{r^3}}, \quad (3)$$

where G is the gravitational constant and $M_\odot$ is the mass of the central star, which in our model we always assume to be a solar mass star. α is a dimensionless quantity which controls the efficiency of angular momentum transport due to viscosity and was introduced by Shakura and Sunyaev (1973). This parametrization reflects the uncertainties about the processes causing viscosity. The energy dissipation due to viscosity is one of the main heat sources in the disk, especially close to the central star. At larger radii, the irradiation by the central star becomes dominant.

The isothermal sound speed is

$$c_s = \sqrt{\frac{k_b T_c}{\mu m_H}}, \quad (4)$$

where k_b is the Boltzmann constant, μ the mean molecular weight and m_H the mass of a hydrogen atom. We assume that all the energy is dissipated close to the midplane of the disk with T_c being the midplane temperature.

Two and even three-dimensional radiative transfer models of circumstellar disks are powerful tools to explore their physical structure and chemical evolution. Very sophisticated models of circumstellar accretion disks were developed in the last years (see Dullemond et al. (2006) for a review). Here, we restrict ourselves to simple analytic models. We explore three different disk models here: Two analytic one-dimensional models and a two-dimensional model. In what follows, we derive the surface density, the temperature and the pressure profiles that result from each model.

2.1.1. Self similar solution

The self-similar solution of the evolution Eq. (1) derived by Lynden-Bell and Pringle (1974) has the form

$$\Sigma(r,t) = \frac{C}{3\pi\nu} A^{-(5/2-\gamma)/(2-\gamma)} \exp\left(-\frac{\tilde{r}^{(2-\gamma)}}{A}\right), \quad (5)$$

where

$$A = \frac{t}{t_s} + 1, \quad (6)$$

$$t_s = \frac{1}{3(2-\gamma)^2}\frac{r_s^2}{\nu}, \quad (7)$$

with $\tilde{r} = r/r_s$, where r_s is the scale length. It gives the transition radius at which the disk's exponential cut-off starts. This self-similar solution holds if the viscosity is given by a power law, $\nu(r) \propto (r/r_s)^\gamma$. We adopt $\gamma = 1$. In the α-prescription, (2) and (4) result in a temperature profile with a fixed power law in r:

$$T_c(r) \propto \tilde{r}^{-1/2}. \quad (8)$$

Then, $\nu(r_s)$ is obtained by (2). C is a constant and can be estimated if the initial disk mass is known.

In principle, the above equation describes a time-independent midplane temperature. Since the surface density evolves in time, we are interested in a time-dependent temperature profile. Hence, we use the above power-law of $T_c \propto \tilde{r}^{-1/2}$ and derive a surface density-dependent normalization of the profile. To estimate the temperature at r_s, we follow Armitage (2010), assuming that the disk surface emits like a black body and the vertical energy flux equals the dissipation rate at the disk midplane and is constant in height. The temperature of the disk at r_s is given by

$$T_c^4(r_s,t) = \frac{27GM_\odot}{64\sigma}\kappa\nu(r_s)\Sigma(r_s,t)^2 r_s^{-3}, \quad (9)$$

where κ is the opacity and σ is the Stefan–Boltzmann constant. The normalization of the temperature profile becomes time dependent. This approach is not fully self-consistent but it provides a simple model. Finally, the pressure P is given by the ideal gas equation of state by:

$$P(r,t) = \frac{\Sigma}{2}\sqrt{\frac{GM_\odot}{r^3}\frac{k_B T_c(r,t)}{m_H \mu}}. \quad (10)$$

2.1.2. Chambers model

A problem that occurs in the self-similar solution model is that the viscosity is constant in time even when the disk cools and the opacity is constant with radius even very close to the star, where all grains will sublimate. Chambers (2009) describes another analytic model which takes into account heating by viscous accretion and radiation from the central star. Three different regimes are introduced: in the innermost disk, where dust grains are sublimated, the opacity follows a power law. Then, for the intermediate part of the disk, a constant opacity is assumed. Viscous heating is the dominant heat source in these regimes. In the third regime, the outer disk, stellar irradiation is the most dominant heat source.

In order to obtain a first approximation, in most parts of the disk we use a constant opacity κ_0. Although more complex, piecewise continuous opacity laws are used widely (Ruden and Pollack, 1991; Stepinski, 1998), there are still uncertainties in the details.

In the intermediate regime, where only viscous heating takes place and the opacity is constant, the disk surface density and temperature are given as a function of time and radius as follows:

$$\Sigma(r,t) = \Sigma_{\rm vis}\left(\frac{r}{s_0}\right)^{-3/5}\left(1+\frac{t}{\tau_{\rm vis}}\right)^{-57/80}, \quad (11)$$

$$T(r,t) = T_{\rm vis}\left(\frac{r}{s_0}\right)^{-9/10}\left(1+\frac{t}{\tau_{\rm vis}}\right)^{-19/40}. \quad (12)$$

$\Sigma_{\rm vis}$ and $T_{\rm vis}$ are the initial surface density and temperature at the outer disk radius. They are given by:

$$\Sigma_{\rm vis} = \frac{7M_{\rm disk}(0)}{10\pi s_0^2}, \quad (13)$$

$$T_{\rm vis} = \left(\frac{27}{64}\frac{\kappa\alpha\gamma k_B}{\sigma\mu m_H}\right)^{1/3}\Sigma_{\rm vis}^{2/3}\Omega_0^{1/3}, \quad (14)$$

where $\gamma \simeq 1.4$ is the adiabatic index, $\Omega_0 = \sqrt{GM_\odot/s_0^3}$ is the Keplerian velocity at the initial disk radius and s_0 is the initial outer radius of the disk, which is the scale length of this model.

The viscous time scale $\tau_{\rm vis}$ is given by

$$\tau_{\rm vis} = \frac{1}{16\pi}\frac{\mu m_H}{\alpha\gamma k_B}\frac{\Omega_0 M_0}{\Sigma_{\rm vis} T_{\rm vis}}. \quad (15)$$

As the surface density of a disk decreases with time, the stellar irradiation will become a dominant heating source and an initially fully viscously heated disk will start to become mostly heated by irradiation in the outer regions, beyond a radius r_t. Then, at $r > r_t$,

$$\Sigma(r,t) = \Sigma_{\rm rad}\left(\frac{r}{s_0}\right)^{-15/14}\left(1+\frac{t}{\tau_{\rm vis}}\right)^{-19/16}, \tag{16}$$

where

$$\Sigma_{\rm rad} = \Sigma_{\rm vis}\left(\frac{T_{\rm vis}}{T_{\rm rad}}\right), \tag{17}$$

with

$$T_{\rm rad} = \left(\frac{4}{7}\right)^{1/4}\left(\frac{T_\odot k_B R_\odot}{GM_\odot \mu m_H}\right)^{1/7}\left(\frac{R_\odot}{s_0}\right)^{3/7} T_\odot. \tag{18}$$

Moreover,

$$T(r,t) = T_{\rm rad}\left(\frac{r}{s_0}\right)^{-3/7}. \tag{19}$$

The radius r_t can be expressed as

$$r_t(t) = s_0\left(\frac{\Sigma_{\rm rad}}{\Sigma_{\rm vis}}\right)^{70/33}\left(1+\frac{t}{\tau_{\rm vis}}\right)^{-133/132}, \tag{20}$$

and this transition radius r_t moves inwards with time. These equations hold if the initial outer edge of the disk is always in the viscous regime, $s_0 < r_t$, which means that irradiation can be initially neglected at the outer disk edge. This is a suitable assumption in our case.

In the regime very close to the star, Chambers (2009) assume that the opacity κ depends on the temperature

$$\kappa = \kappa_0\left(\frac{T}{T_e}\right)^n, \tag{21}$$

where T_e is a transition temperature which defines the border between the inner most region and the intermediate regime. Stepinski (1998) recommended $T_e = 1380$ K and $n = -14$. The surface density and temperature change with time as follows:

$$\Sigma(r,t) = \Sigma_{\rm evap}\left(\frac{r}{s_0}\right)^{-24/19}\left(1+\frac{t}{\tau_{\rm vis}}\right)^{-17/16}, \tag{22}$$

with

$$\Sigma_{\rm evap} = \Sigma_{\rm vis}\left(\frac{T_{\rm vis}}{T_e}\right)^{14/19}, \tag{23}$$

and

$$T(r,t) = T_{\rm vis}^{5/19} T_e^{14/19}\left(\frac{r}{s_0}\right)^{-9/38}\left(1+\frac{t}{\tau_{\rm vis}}\right)^{-1/8}. \tag{24}$$

The transition radius between the inner and the outer viscous regimes has the form

$$r_e(t) = s_0\left(\frac{\Sigma_{\rm evap}}{\Sigma_{\rm vis}}\right)^{95/63}\left(1+\frac{t}{\tau_{\rm vis}}\right)^{-19/36} \tag{25}$$

and moves inwards with time.

2.1.3. Two-dimensional model

The analytic models are given by radius-dependent quantities and basically all quantities in the vertical direction are represented by their central or averaged values. Following Papaloizou et al. (1999), Hersant et al. (2001) and Huré (2000), we estimate the pressure and temperature profiles of the disk in a more sophisticated way in a two-dimensional model: in radial direction r and in vertical direction z. For each radius, a set of differential equations describes the vertical structure of the disk, where we ignore self-gravity. The structure is described (Armitage, 2010) by hydrostatic equilibrium

$$\frac{1}{\rho}\frac{dP}{dz} = -\Omega^2 z, \tag{26}$$

by the vertical variation of the flux $F(z)$

$$\frac{dF}{dz} = \frac{9}{4}\rho(z)\nu(z)\Omega^2, \tag{27}$$

and by the radiative heat flux in the diffusion approximation for radiative transfer,

$$\frac{dT}{dz} = -\frac{3\kappa\rho(z)}{16\sigma T(z)^3}F(z). \tag{28}$$

In addition, there is an equation of state

$$P(z) = \frac{\rho(z)k_B T}{\mu m_H} + \frac{4\sigma}{3c}T(z)^4. \tag{29}$$

The second term is the radiation pressure which is usually very small and we ignore it. Turbulent pressure can be ignored if $\alpha \ll 1$, Huré (2000).

The opacity κ is in general a function of P and T. In this model we use an opacity model presented by Ruden and Pollack (1991), which is a piecewise, continuous function of temperature. More sophisticated models add an additional atmosphere-like optically thin layer above the optically thick disk (Huré, 2000), which we neglect.

Papaloizou et al. (1999) used the above set of equations to estimate the vertical structure of protoplanetary disks. The pressure $P(H)$ is, with a boundary optical depth of $\tau(H) = 2/3$, given by

$$P(H) = \frac{2H\Omega^2}{3\kappa}. \tag{30}$$

Papaloizou et al. (1999) pointed out that the results do not depend on the exact choice of $\tau(H)$ if $\tau(H) \ll 1$. At least the midplane pressure does not depend sensitively on this value, although this need not be the case for values at other scale-heights.

The accretion rate $\dot{M}_{\rm acc}(t)$ in the disk decreases in time according to (Drouart et al., 1999),

$$\dot{M}_{\rm acc}(t) = \dot{M}_{\rm acc}(0)\left(1+\frac{t}{t_0}\right)^{-3/2}, \tag{31}$$

where

$$t_0 = \left[\frac{R_D^2}{3\nu(R_D)}\right]_{t=0}. \tag{32}$$

R_D is the time dependent outer radius of the disk. The initial accretion rate is assumed to be $\dot{M}_{\rm acc}(0) = 5\times10^{-6} M_\odot\ {\rm yr}^{-1}$.

Each annulus of the disk can be calculated independently. To solve the set of Eqs. (26)–(29), initially, a scale height H is guessed. Calculations start at the top of the disk located at height H, where the boundary conditions are given by the optical depth $\tau(H) = 2/3$, and

$$F(H) = \frac{3}{8\pi}\dot{M}_{\rm acc}(t)\Omega^2. \tag{33}$$

In the midplane, the energy flux along the vertical direction must be $F(0) = 0$ due to disk symmetry. If the estimated flux $F_{\rm est}(0)$ in the first calculation is not sufficiently small, that is $|F_{\rm est}(0)/F(H)| < 10^{-4}$, a new H is chosen, estimated through a Newton–Raphson method. Given the initial disk radius R_D, α-parameter and the initial accretion rate, we can thus calculate density, temperature and pressure profiles at any time t. Huré and Galliano (2001)

reports of a 30% maximum deviation on the midplane temperature comparing one-dimensional and two-dimensional models.

2.2. Chemistry

To estimate the chemical composition of solid material in the protoplanetary disk midplane, equilibrium condensation of the gas is assumed. In a closed system of elements, condensation calculation models describe the equilibrium distribution of the elements between coexisting phases. For decades, condensation calculations have been used to describe chemistry in astrophysical problems, for example the formation of the Earth (Latimer, 1950). Equilibrium condensation provides a basic framework to predict and interpret bulk chemical composition of meteorites and planets (Ebel, 2006; Davis, 2006; Righter et al., 2006). Hence, condensation calculations are a good starting point for our study. Following previous studies (Pasek et al., 2005; Bond et al., 2010a), we use the commercial HSC Chemistry software package to perform the equilibrium condensation calculations. The equilibrium chemical composition of a system of elements is found by iteratively minimizing it's Gibbs free energy, which is done in the GIBBS equilibrium solver. The algorithm is further described in White et al. (1958). Over one hundred different gaseous and solid species are taken into account, see Bond et al. (2010a). The initial parameters for the HSC Chemistry equilibrium composition calculation module are the amount of each pure gaseous element in kmol. It is assumed that no other species are present initially. In addition, the solids that are present when equilibrium is reached are also specified. Starting from a gas of solar composition (Asplund et al., 2005), see Table 1, and under the assumption of chemical equilibrium, we calculate which distribution of compounds is most thermodynamically stable at midplane pressure $P(r,t)$ and temperature $T(r,t)$ given by the disk models. We do the calculations in a radial range of $r = [0.25\ \text{AU}, 4\ \text{AU}]$ with a step size of $\Delta r = 0.05$ AU.

The output is the amount of each compound in kmol. We are interested in the relative abundances of the elements in the condensates. Therefore, we weight the abundance of a certain species with the number of atoms of the element included in that certain molecule and its atom weight. The relative abundance is expressed as wt.% which means *weight percent(age)*:

$$\text{wt.\% of } X = \frac{\sum_A m_X X_A n_A(r)}{\sum_A \sum_Y m_Y Y_A n_A(r)}, \tag{34}$$

where n_A is the total abundance of species A located at radius r and X_A gives the number of X-atoms in species A. m_X is the atom weight of the element. The numerator is a summation over all solid species that contain the element X. The denominator is a summation over the weight of all atoms Y in all species A in the solid phase at radius r.

Table 1
Relative element abundances in the present-day solar photosphere from Asplund et al. (2005). The abundance is given as the exponent to the base 10. Hence, H has the highest abundance in the solar photosphere, its relative abundance is 10^{12}. If this value is used as input for the HSC Chemistry module, it can be interpreted as the number of moles of H, whereas the normalization is arbitrary and can be neglected.

Element	Abundance	Element	Abundance
H	12.00	Si	7.51
He	10.93	P	5.36
C	8.39	S	7.14
N	7.78	Ca	6.31
O	8.66	Ti	4.90
Na	6.17	Cr	5.64
Mg	7.53	Fe	7.45
Al	6.37	Ni	6.23

In the following, we often order the involved elements according to their volatility. A good measure of the volatility is the temperature at a given pressure at which 50% of the original element has condensed (Lodders, 2003). In the case of O (and C), this temperature is lower than any temperature reached in our disk models, since most O condenses in water ice. Therefore, its volatility is measured according to the condensation of the higher temperature condensates.

2.3. Dynamical simulations

Morishima et al. (2010) carried out 64 simulations which describe the collisional growth of planetesimals and the subsequent long-term evolution and stability of the resulting planetary systems. The simulations explore sensitivity to the initial conditions, including the initial mass and radial distribution of planetesimals, the timescale for the dissipation of the solar nebula and different orbits of Jupiter and Saturn. The simulations start with 2000 equal-mass particles placed between 0.5 and 4 AU embedded in a gas disk. The planetesimal disk has an initial mass $m_{\text{planetesimals}}$ of 5 $m_\oplus$ or 10 $m_\oplus$. Since we use this mass very often in this report, we introduce m5 for the case of a dynamical simulations with $m_{\text{planetesimals}} = 5\ m_\oplus$ or m10 for the case with $10\ m_\oplus$ The surface density Σ of solids and of the initial gas disk depends on the radius, $\Sigma \propto r^{-p}$, where p is 1 or 2. The gas disk dissipates uniformly in space and exponentially in time with a gas dissipation time scale $\tau_{\text{gas}} = 1,\ 2,\ 3$ or 5 Myr. After time τ_{gas} from the beginning of the simulation, when a significant part of the gas disk has disappeared, Jupiter and Saturn are introduced on their orbits. Several orbits are tested: CJS are Circular orbits according to the initial conditions used in the Nice model (Tsiganis et al., 2005), EJS place Jupiter and Saturn on their current orbits, EEJS are the same as EJS but higher eccentricity for Jupiter, CJSECC are the same as CJS but higher eccentricities for Jupiter and Saturn. Each simulation represents one unique combination of initial conditions. See Tables 1 and 2 in Morishima et al. (2010) for further details.

The simulations in Morishima et al. (2010) were constructed in an attempt to reproduce the global characteristics of the terrestrial planets in the Solar System. The constraints which were used to compare the simulation results with the observations are the spatial mass distribution, the deviation from circular and coplanar orbits and the timing of the Moon-forming impact. In general, 1–5 planets are formed in the terrestrial region. An observed trade-off is that if a similar radial mass concentration is achieved, the Moon-forming impact occurs too early. The dependence on the initial conditions becomes manifest for example in the fact that small gas dissipation timescales ($\tau = 1$—2 Myr) are needed to avoid a significant depletion of solid material due to migration.

Other simulations (O'Brien et al., 2006) often consist of initially roughly Mars-mass embryos embedded in a planetesimal disk with no mutual interaction between the planetesimals. The initial mass of the fully interacting particles in the Morishima et al. (2010) simulations is smaller than the Moon mass. Hence, Mercury analogs concerning its mass and semi-major axis can form in those simulations in contrast to other studies. Nevertheless, these calculations assume perfect accretion and no mass loss due to giant impacts – such an event is most probably responsible for striping off Mercury's initial crust and mantle (see Benz et al. (2007) for a review).

The Minimum Mass Solar Nebula (Hayashi, 1981) has a gas density profile $\Sigma \propto r^{-p}$ with $p = 1.5$. The steady-state solution of (1) predicts $p = 1$. Observations suggest a flatter profile with $p < 1$ (Kitamura et al., 2009; Andrews et al., 2009). Therefore, a gas density profile with $p = 2$ is extremely steep in the context of circumstellar disks and we focus on the dynamical simulations with $p = 1$. The initial surface density of solids at 1 AU is $\Sigma_{0,\text{solid}} = 6.1\ \text{g cm}^{-2}$ in the case of m5 or $\Sigma_{0,\text{solid}} = 12.2\ \text{g cm}^{-2}$ for

m10. The gas density at 1 AU is $\Sigma_{0,\mathrm{gas}} = 2000\ \mathrm{g\ cm^{-2}}$ for both m5 and m10.

2.4. Transition from disk models to dynamical simulations

In order to combine the disk models and the dynamical simulations, we have to choose a suitable transition criterion. Since our goal is to study the evolution of the solid material in the dynamical simulations, we define the solid surface density at 1 AU as the normalization condition. This means that the surface density of solids Σ_{solid} predicted by the disk model at 1 AU at transition time $t = t_{\mathrm{trans}}$ has to reproduce the initial solid surface density of the dynamical simulations at 1 AU, $\Sigma_{0,\mathrm{solid}}$. For example,

$$\Sigma_{\mathrm{solid}}(1\ \mathrm{AU}, t_{\mathrm{disk}}) = \Sigma_{0,\mathrm{solid}} = 6.1\ \mathrm{g\ cm^{-2}}, \tag{35}$$

in the case of the m5 simulations. Actually, the disk models give the surface density of gas Σ_{gas} and we assume a gas-to-dust ratio of

$$\frac{\Sigma_{\mathrm{gas}}(r,t)}{\Sigma_{\mathrm{solid}}(r,t)} = 100, \tag{36}$$

which is roughly the gas-to-dust ratio of the ISM. This value is smaller than the gas-to-dust ratio in Morishima et al. (2010);

$$\frac{\Sigma_{0,\mathrm{gas}}}{\Sigma_{0,\mathrm{solid}}} = 329, \tag{37}$$

for m5 simulations, and half that for m10. These ratios are rather high and can only be achieved with a very high solid mass loss rate combined with a low gas dissipation rate. Since we model the solids in this disk, the ISM ratio seems to be a better choice.

The initial total disk mass in the dynamical simulations is

$$m_{\mathrm{tot}} \approx m_{\mathrm{planetesimals}} \frac{\Sigma_{0,\mathrm{gas}}}{\Sigma_{0,\mathrm{solid}}} = 1650\ m_{\oplus} \cong 0.0055 M_{\odot}, \tag{38}$$

for m5 and the m10 simulations. Therefore, the initial mass of $0.1 M_{\odot}$ in our disk models is a justified assumption since the disk loses mass with time. Disk model masses at $t = t_{\mathrm{trans}}$ are shown in Table 2.

A disk model is described by the initial mass, time t, the α-parameter and the disk scale radius r_s. For all models we adopt the following initial constraints: the initial disk mass $M_{\mathrm{disk}}(t=0) = 0.1 M_{\odot}$ and $\alpha = 0.009$, Hersant et al. (2001). The models use different opacity laws. By default, the constant opacity is $\kappa_0 = 3\mathrm{cm^2\ g^{-1}}$. We adopt a radial scale length $r_s = 10$ AU. We are only interested in profiles that reproduce the surface density at 1 AU. The choice of the scale radius has no significant effect on the results: In the self-similar solution model, the effect of the scale radius r_s on the surface density distribution can be compensated by a suitable choice of t_{trans} and C. It has a minor effect on temperature and pressure.

In the Chambers model, the initial disk radius and the time t are free parameters and they are degenerate in a similar way as in the self-similar solution model. In the two-dimensional model, the accretion rate is a function of the initial outer disk radius, time and initial accretion rate (see Eq. (31)). Since the surface density in this model is controlled by the accretion rate, the disk model is also degenerate in the parameters R_D and t_{trans}. Actually, the choice of the scale radius is not completely free: In the case of a very large initial scale radius ($r_s \sim 100$ AU), the initial disk mass might be too small to reproduce the surface density at 1 AU for any $t > 0$. For $r_s = 10$ AU, there is a $t_{\mathrm{trans}} > 0$ such that $\Sigma(1\ \mathrm{AU}, t_{\mathrm{disk}}) = \Sigma_{0,\mathrm{solid}}$ for all models. The same holds for the choice of the initial accretion rate in the two-dimensional model. If the initial accretion rate is too small, it will never result in a suitable surface density. Table 2 lists the model-dependent transition times which can differ by orders of magnitude.

Note that in Bond et al. (2010a), the transition time is left as a free parameter. Hence, a better fit to the observed planetary abundance values is possible, but this approach is not as self-consistent as ours.

Table 2
The adopted transition time and the resulting disk mass for the different disk models and solid disk masses. r_{in} is the innermost semi-major axis at which solids condense at t_{trans}. The model parameters are $\alpha = 0.009$ and $r_s = 10$ AU and the initial disk mass is $m_{\mathrm{disk}}(0) = 0.1\ M_{\odot}$.

Disk model	Disk mass	t_{trans} (10^4 yr)	M_{disk} ($10^{-2}\ m_{\odot}$)	r_{in} (AU)
Self-similar solution	$m_T = 5\ m_{\oplus}$	3.7	3.5	0.50
Self-similar solution	$m_T = 10\ m_{\oplus}$	2.1	4.4	0.75
Chambers	$m_T = 5\ m_{\oplus}$	5.0	5.0	0.30
Chambers	$m_T = 10\ m_{\oplus}$	1.9	6.0	0.50
Two-dimensional	$m_T = 5\ m_{\oplus}$	25.0	1.2	0.35
Two-dimensional	$m_T = 10\ m_{\oplus}$	41.0	1.8	0.65

3. Results

This section presents the different disk models that are normalized to the initial conditions of the dynamical simulations. The resulting chemical abundance profiles are shown briefly. Then, the source regions of the planets are studied. Combining the chemistry and the dynamics results in the bulk chemical composition of the planets which are studied in detail. Two elements which are very important for the emergence and evolution of life, H and C, are briefly highlighted at the end of the section.

3.1. Initial disk profiles

Surface density profiles and the corresponding pressure and temperature profiles at t_{trans} for the different models are plotted in Fig. 1. The slope in the temperature and pressure profiles of the models are very similar. At radial distance $r > 1$ AU, the different opacity models change the slope of the temperature profile. The slope of the dynamical simulation surface density ($\Sigma \propto r^{-1}$) is only reproduced in the self-similar model with the other models yielding flatter profiles. The high solid surface density of the m10 simulations results in disk models that predict a high temperature and pressure profiles relative to the m5 disk. The Chambers model gives the highest overall temperatures, especially at larger radii. In general, the two-dimensional model gives the lowest temperature. These differences allow us to explore a range of different temperature and pressure profiles in our study.

Comparison with disk profiles obtained in previous studies is difficult since our initial assumptions vary. A brief comparison with e.g. the sophisticated study by D'Alessio et al. (1998) reveal that our profiles represent a reasonable approximation to more complex models.

3.2. Elemental abundance profiles

Fig. 2 shows the elemental abundances in solids as a function of radius in the case of the two-dimensional model and a disk of mass m5 and the Chambers disk model in the case m10 as extrema. The two-dimensional model is a good case study to explain some of the main features since this model gives the largest numbers of solids along the disk: at 1.3 AU, troilite (FeS), at 2 AU, fayalite (Fe_2SiO_4) and around 3 AU iron oxide (Fe_3O_4) all condense out of the solar nebula. Serpentine ($Mg_3Si_2O_5(OH)_4$) and aluminum oxide (Al_2O_3) follow at 3.1 AU. Since all models have similar temperature profile slopes at $r > 1$ AU but predict different temperature normalizations, the location of condensation shifts to larger radii in the case

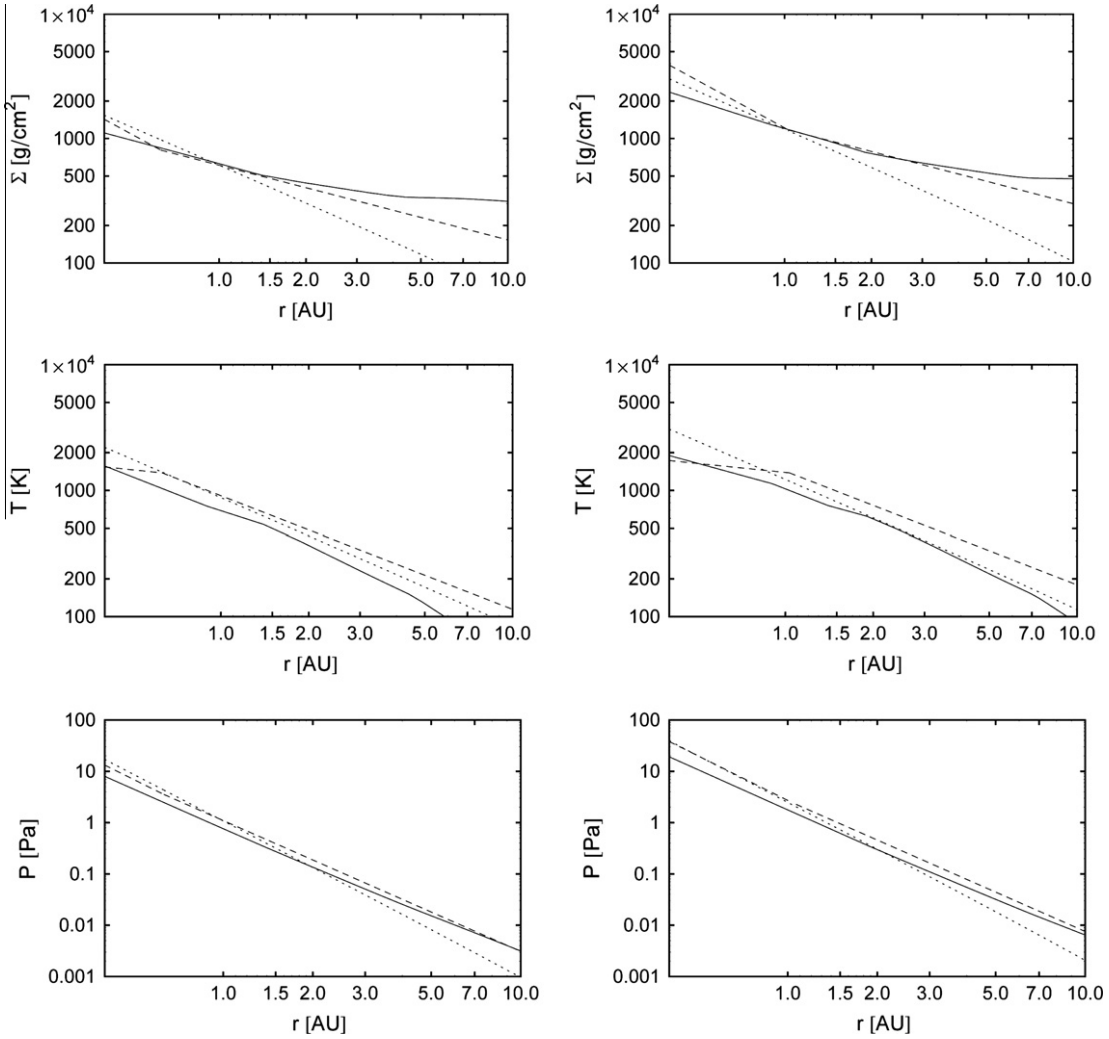

Fig. 1. From top to bottom: surface density profiles, temperature profiles and pressure profiles in the midplane, the m5 disk in the left column, the m10 disk in the right column at transition time. Each plot shows the two-dimensional model (solid line), the Chambers model (dashed line) and the self-similar solution model (dotted line). The adopted model parameters are given in Table 2.

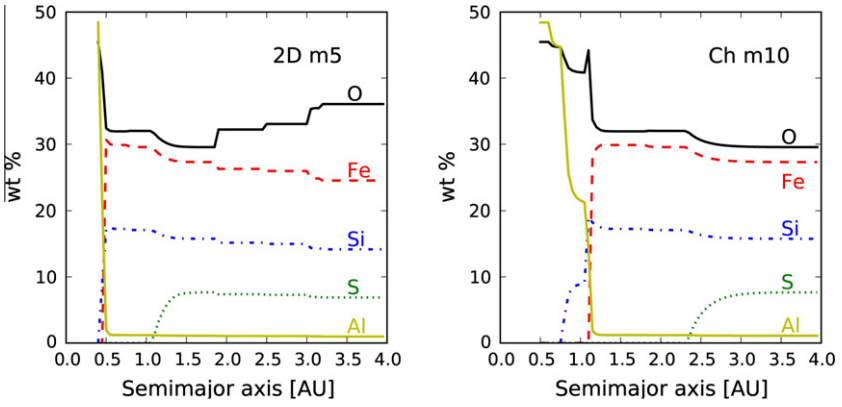

Fig. 2. The weight percentage chemical abundance of some of the most frequent elements in solid species as a function of distance from the star. The left panel gives the two-dimensional model (2D) in the case m5, and the right panel gives the Chambers disk model (Ch) in the case m10. The elements are from top to bottom at 4 AU: O, Fe, Si, S and Al. For example: in the 2D case, a solid particle condensing at 1 AU is composed by weight of 30% O, 30% Fe, 20% Si and 20% is comprised of all other elements.

of colder disks. In the case of the Chambers model in Fig. 2, the condensation of FeS is the only feature beyond ~ 1 AU beside the occurrence of Fe, magnesium silicates and other species around 1 AU.

We point out that the total amounts of solid material changes completely with semi-major axis. Hence, traces of refractory solids that condense close to the star can dramatically change the percentage weights of the different elements. We see such features

in the profile close to the star. If the temperature and pressure are high enough ($T \gtrsim 1340$ K, $p \gtrsim 10^{-5}$ bar) so that even refractory solids would sublimate, the estimation of the elemental abundance breaks down. Inside this radius $r_{\rm in}$, listed in Table 2, no condensates will form. The self-similar solution model and the two-dimensional model in the m10 disk fail to provide condensates across the whole annulus $a = [0.5\ \text{AU}, 4\ \text{AU}]$ where planetesimals in the dynamical simulations are initially located.

In general, close to the star, only refractory species condense out of the solar nebula. The abundance of more volatile elements increases with distance from the star. Hence, a hot disk model generates a disk with a relatively high abundance of refractory elements. On the other hand, cooler disks provide higher abundances of volatile elements.

3.3. Source regions in the N-body simulations

In the dynamical simulations, the orbits of the planetesimals and protoplanets change due to various mechanisms, such as close encounters or secular resonances. Gas drag damps the eccentricities of the smallest planetesimals and moves them inwards. If they do not grow fast enough and do not merge with other bodies, they do not decouple from the gas. Hence, the planetesimals at small initial semi-major axis fall into the star. On the other hand, larger bodies start to migrate inward due to type I migration as long as the gas disk has a significant surface density. The gas dissipation time scale is of the order of few Myr. As a result, most of the material initially located inside 1 AU is lost in almost every simulation.

The region of initial locations of the planetesimals that end up embodied in a final planet is what we call the planet's source region. In order to study the source regions of individual planets, we group planets first according to their semi-major axis. We divide the planets in four groups (a, b, c and d, ordered by increasing semi-major axis), each containing the same total mass (and thus the same number of initial planetesimals (see Fig. 3)). To study the source region as a function of the planet's final mass in each of these groups, we divide each group a, b, c and d in two sub-groups containing the same number of initial planetesimals: a group of high mass planets and a group of low mass planets. With the statistics of four orbital groups and 4×2 mass sub-groups, the dependence of the source region on the planet's semi-major axis and final mass can be studied. A study of the dependence on the simulation initial conditions would be desirable but the number of simulations is too small. A planet's source region is characterized by the cumulative distribution function of planetesimals as a function of radius from which it is built. We compare the source regions of any two groups using the two-sided Kolmogorov–Smirnov-test. The estimated p-value gives the probability that the two distributions of the initial planetesimal locations are drawn from the same parent distribution. If p is small, it is unlikely that the source regions coincide.

First, we explore how the source regions differ depending on the final orbital radii of the planets. For each planet in a given orbital group, we gather the initial positions of its building blocks. Fig. 3 gives the histograms of the locations of all the initial planetesimals that end up within the planets of the four different orbital groups. Planets close to the star have a wide and flat source region (group a), they form from material that was also closer to the star initially. Planets at larger distances (groups b, c and d) all have steeper and narrower source regions. For these groups, the source regions arise entirely from beyond a radial distance of 1 AU. This is mainly because almost all material initially within this location ends up in the star. For example, the black histogram (group a) indicates that more than 200 planetesimals ($\approx 5\%$ of the total mass of the group) initially located within 1.0 AU end up in the planets in group a, whereas according to the green histogram, not more than 10 planetesimals end up in planets of group d. KS-tests reveal that the groups do not share the same initial distribution as all p-values are below 10^{-6}. This means that we should not expect that planets at different semi-major axes have the same source region. The inner most group a stands out most significantly from the others (see Fig. 3).

Next, we study how the source regions differ depending on the final mass of the planets. The planets in each orbital group are split into two sub groups according to their mass. The KS-test shows that the sub-group containing low mass planets and the sub-group containing high mass planets do not have the same parent distribution in all of the orbital groups. A trend is visible: low mass planets have a source region that is steeper and narrower, whereas massive planets have a wider and flatter source region (see Fig. 3).

Often, the planets are located much closer to the star than the average radii of the initial planetesimals from which they were

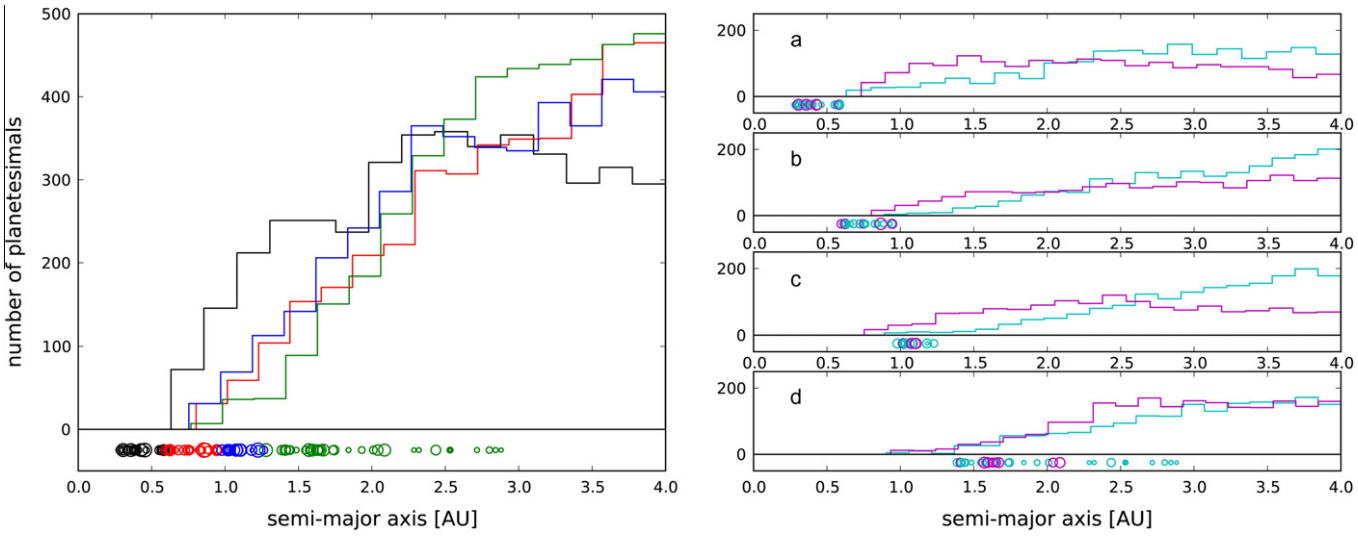

Fig. 3. Left: the initial distribution of all planetesimals that end up in one of the planets in four orbital groups indicated by different colors (a: black, b: red, c: blue, d: green). The lower panel shows the distribution of all planets in radial distance, their diameter represents their masses. Right: the distribution of the planetesimals of the small mass (cyan) and high mass (magenta) sub-groups in the four orbital groups a–d. The lower panel in each histogram gives the distribution of the orbital group that is considered. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

built. This is due to orbital migration. Most of the bulk composition of the planets is contributed by the solids that condense out of the solar nebula at 1 AU$< r <$ 4 AU, except some of the very close-in planets.

3.4. Combining disk chemistry and N-body simulations

In the next step, the calculations of the chemical abundance in the disk model and the dynamical simulations are combined. According to our transition criterion, the disk model dependent transition time $t_{\rm disk}$ estimated in Section 3.1 is the time in the disk model evolution at which the chemical composition of solids in the disk represents the composition of the planetesimals at the beginning of the dynamical simulations. The initial radial distribution of the planetesimals that end up in a given planet gives the final bulk chemical composition of this planet. The element abundance profiles in the disk and the diversity in the source regions suggest that different planets have different bulk chemical compositions. All planetesimals from the dynamical simulations that contribute to the mass of any planet originate from beyond the condensation line $r_{\rm in}$ at $t_{\rm trans}$. There the composition of all planets can be fully determined. The element abundance in the planets for each disk model is shown in the online Supplementary data.

The simulations by Morishima et al. (2010) are designed to reproduce the inner planets of the Solar System, thus, we now compare the element bulk abundance of the simulated planets with the Solar System terrestrial planets.

A normalized abundance $N_{\rm X}$ of element X relative to Si can be obtained from (34) via:

$$N_{\rm X,tp} = \frac{\left(\frac{\text{wt.\% of X}}{\text{wt.\% of Si}}\right)_{\rm sp}}{\left(\frac{\text{wt.\% of X}}{\text{wt.\% of Si}}\right)_{\rm tp}} \tag{39}$$

where the wt.% is based on simulated planets (subscript sp) or literature values of terrestrial planets (subscript tp) in our Solar System. The bulk chemical compositions were taken from Morgan and Anders (1980), Kargel and Lewis (1993), and Lodders and Fegley (1997), see Table 3. One should keep in mind that these values are partly based themselves on simulations and that uncertainties lead to absolute errors up to the order of 25% on these abundances (Bond et al., 2010a). These uncertainties are discussed in the caveats Section 4.1.

Next, we search for a simulation outcome that best reproduces the composition of the inner Solar System planets. We choose a planetary system including four planets out of the m10 simulations and a similar system out of the m5 simulations. Since all planets have to form in the same simulation and the number of simulations is limited, we focus on systems that match the constraints of our Solar System best according to Morishima et al. (2010), keeping in mind that differences in mass and semi-major axis have an effect on the source region. EJS 2-1-5 (short hand for $\tau_{\rm gas} = 2$ Myr, $p = 1$, m5 (see Section 2.3)) and EJS 3–1-10 are two systems that fulfill many of the constraints, although the masses of the planets are very poorly reproduced: Venus and Earth are too small and the Mars and Mercury analogs are too massive. The source region of close in planets is sensitive to the planets mass and the large Mercury mass in the EJS 2-1-5 simulation might affect the result. The large Mars mass is a general problem that is often observed in dynamical simulations that reproduce the inner Solar System, recently discussed in Walsh et al. (2011). The masses and semi major axis of the chosen simulated planets are shown in Table 4.

In Fig. 4, the normalized abundance $N_{\rm X,tp}$ is shown for each element and all Solar System planets. Both systems fail in reproducing in detail the bulk chemical composition of the Solar System. The most refractory elements (Al, Ti and Ca) are underestimated in Mercury, Venus and Earth analogs. There are two reasons for this: Gas drag and type-I migration leads to the loss of refractory rich planetesimals that form in the inner part of the disk. Hence, they cannot be embedded in the planets. On the other hand, if the temperature of the disk is high enough so that refractory solids dominate the solid abundance not only in the very center of the disk, more planetesimals are dominated by refractory material. The Chambers model in the m10 case predicts the highest overall temperature and the normalized abundances of refractory material for the Earth and Mars analogs are significantly increased. The most volatile element abundances (Na and S) are overestimated. This discrepancy is higher if the disk is too cold. In this case volatile rich planetesimals dominate the disk and the normalized bulk abundance becomes too high in the case of all four planets. Except in the case of S, the composition of Mars is reproduced by most of the disk models. On the other hand, Mercury is not very well reproduced in any of the models.

Note that the differences in bulk composition of the planets we are considering is small. In both simulations, the abundances differ by up to 50% in the case of refractory elements if the Chambers-model is used. All the other element abundances nearly coincide. The fact that the variation in bulk chemical composition in the simulated planets is in general smaller than in the Solar System rocky planets is also demonstrated in Fig. 5.

Terrestrial planet formation and especially the chaotic growth phase when many giant impacts take place is a stochastic process. This might result in the formation of unique planets that reproduce the bulk chemical composition of certain Solar System planets very well, but to reproduce the bulk composition of all Solar System planets in a single simulation might be a very rare situation. Hence, we study the distribution of the element abundance of all planets at once to get an overview as to how extreme the composition of

Table 3
Planetary abundances in wt.% or weight ppm.

	Mercury[a]	Venus[a]	Earth[b]	Mars[c]
Fe (%)	64.47	31.17	32.04	27.24
O (%)	14.44	30.9	31.67	33.75
Mg (%)	6.5	14.54	14.8	14.16
Al (%)	1.08	1.48	1.43	1.21
Si (%)	7.05	15.82	14.59	16.83
Ni (%)	3.66	1.77	1.72	1.58
S (%)	0.24	1.62	0.89	2.2
Ca (%)	1.18	1.61	1.6	1.33
C (ppm)	5.1	468	44	2960
N (ppm)	0.046	4.3	0.59	180
Na (ppm)	200	1390	2450	5770
P (ppm)	390	1860	1200	1100
Ti (ppm)	630	850	800	650
Cr (ppm)	7180	4060	3400	3680

[a] Morgan and Anders (1980).
[b] Kargel and Lewis (1993).
[c] Lodders and Fegley (1997).

Table 4
Parameters of simulated planetary systems that represent the Solar System. a is the semi-major axis of the planet. The Solar System values are taken from Kargel and Lewis (1993).

	Mercury	Venus	Earth	Mars
EJS 2-1-5				
Mass ($m_\oplus$)	0.52	0.45	0.79	0.51
a (AU)	0.31	0.59	0.94	1.39
EJS 3-1-10				
Mass ($m_\oplus$)	0.09	0.23	0.34	0.27
a (AU)	0.67	0.89	1.26	1.54
Solar System				
Mass ($m_\oplus$)	0.06	0.81	1.00	0.11
a (AU)	0.38	0.72	1.00	1.52

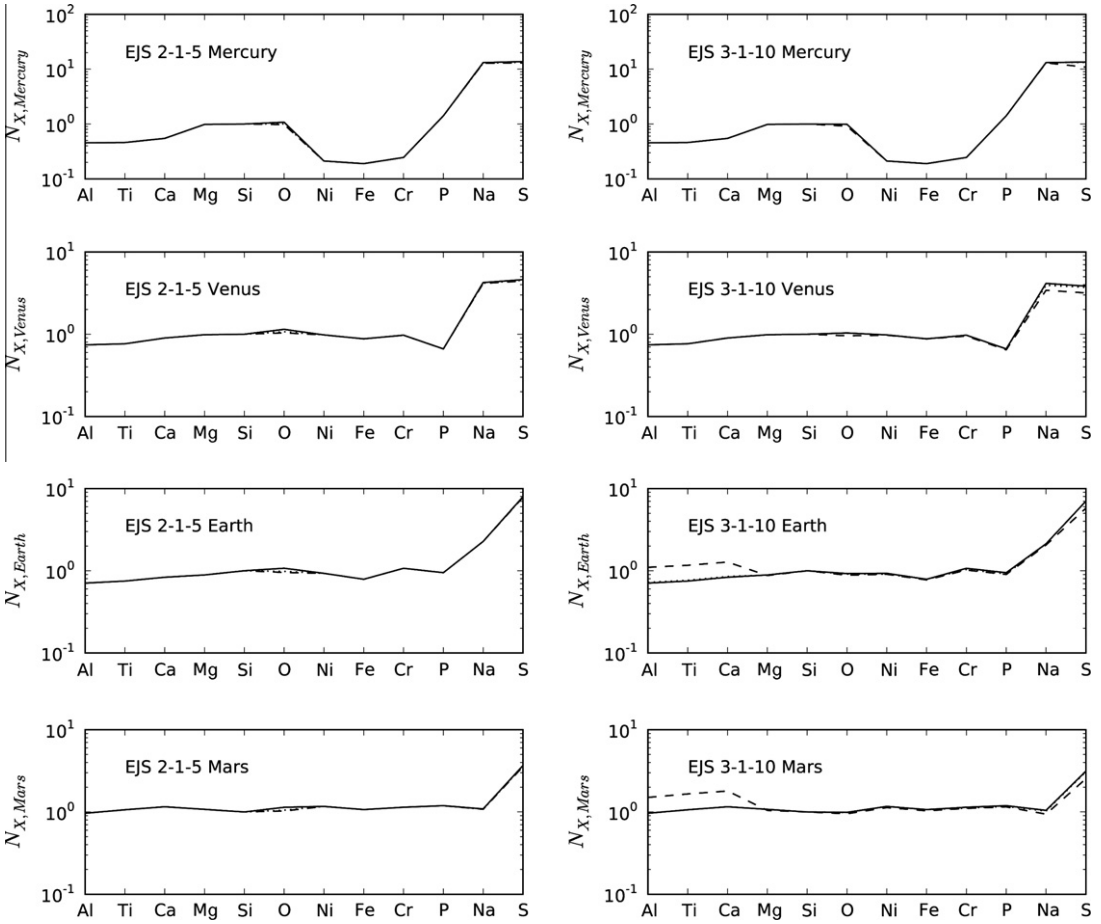

Fig. 4. The normalized abundance of elements for Mercury, Venus, Earth and Mars analogs, from bottom to top. The elements are listed in order of increasing volatility from left to right. Two-dimensional model (solid line), Chambers model (dashed line) and self similar solution model (dotted line) disks are taken into account. Left column: EJS 2-1-5, right column: EJS 3-1-10.

planets in those simulations can be. As source regions of different planets can differ significantly, we expect variations in the element abundances depending on the underlying disk model. In Fig. 5, the distribution of the elemental bulk chemical compositions of planets is shown for each element in the case of the most extreme disk models normalized to the Earth values. One model is the two-dimensional model in the case m5. This is the model that gives the lowest temperature across the disk and a lot of species condense across the planetesimal annulus. Thus, it results the most inhomogeneous element abundance profile. The other model is the Chambers model in the m10 case where the resulting element abundance profiles are the most featureless since this disk has the highest temperature. Although these models are applied one two different samples of planets, comparing their element abundances reveals the most extreme variations that are possible in our approach.

In the two-dimensional model, the relative abundances of the simulated planets are mostly located in a narrow range. In the case of refractory elements, the abundances are a little lower than the Earth value and they agree with Mars. There is a variation in O, especially for small semi-major axes. In the case of the volatile elements Na and S, the bulk chemical composition is very high relative to the Earth values, up to a factor of 9 in the case of sulfur. Some planets with small semi-major axis tend to have a smaller abundance of these volatile elements.

In the Chambers model, there is a variation in the relative abundances, up to a factor of 5 in the case of sulfur. Concerning refractory elements, some planets still group along a narrow range at the same position as in the two-dimensional model. They form a lower bound of the distribution. Some planets have much higher element bulk abundances and agree with the relative abundances of Mercury. Beside some outliers, these planets have small semi-major axes. The volatile element abundances are reproduced in a similar manner than before, although the variation towards lower relative abundances is a little larger.

The bulk chemical compositions are much more sensitive to the different source regions than in the cooler two-dimensional model. This can be explained using the example of sulfur: in the two-dimensional model FeS condenses around 1 AU, hence most planets incorporate planetesimals with a similar weight percentage of sulfur. The few outliers result from the difference in the source regions. In the Chambers model, condensed sulfur is only present outside 2.5 AU. Hence, depending on the source region, a larger variation in the bulk composition is possible.

Another example is given by the refractory element Al. In the hot disk, Al is the dominant element inside the condensation radius of more volatile solids at around 1.3 AU (see Fig. 2). Hence, planetesimals that are locate initially inside this radius have a wt.% that is higher than any planetesimals that form in the colder disk. This results in the higher variation.

In both models, the differences between the simulated planets and the Solar System planets are significant, especially in the volatile regime and in the case of Mercury.

Geochemical ratios allow us to quantify the depletion in volatile elements (e.g. Na/Si) with respect to the enrichment in refractory elements (e.g. Al/Mg). The ratios in the solar photosphere (Table 1)

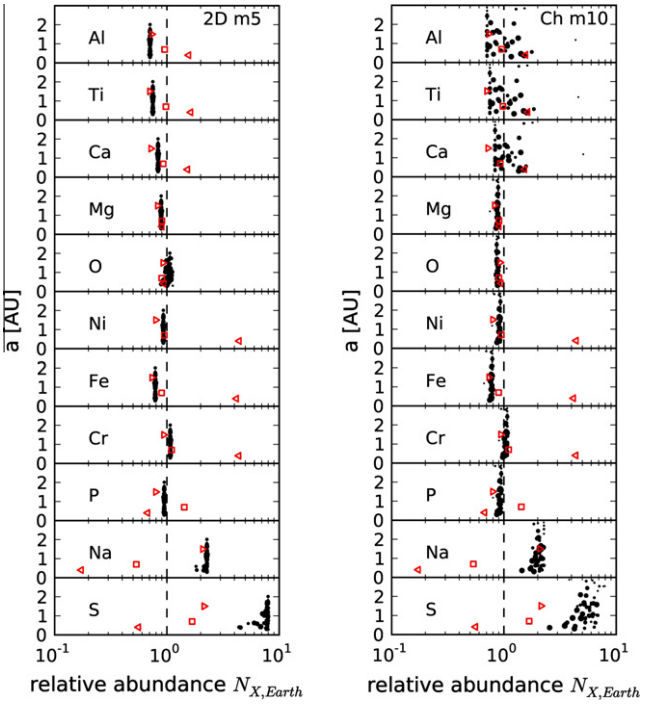

Fig. 5. The element abundance of all planets. The ordering of the elements increase in volatility from top to bottom. The horizontal axis gives the abundance of element X of a planet relative to the abundance of Si, see (39) normalized by Earth values. The vertical axis gives the semi-major axis of the planet. The diameter of the black dots represents the mass of the simulated planets. The red markers indicate location of Mercury (left triangle), Venus (square) and Mars (right triangle), the dashed line highlights the Earth abundance. The left column represents m5 with a two-dimensional disk model (2D), this is the model that predicts the lowest disk temperature. The right column represents m10 with a Chambers disk model (Ch), this is the model that predicts the highest disk temperature. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

are Na/Si = 0.045 and Al/Si = 0.072. For the Earth, according to Table 3, Na/Si = 0.017 and Al/Si = 0.098. This reflect the volatile depletion and refractory enrichment in the Earth relative to Solar System bulk composition (Palme et al., 2003; Rubie et al., 2011). Fig. 5, which gives the ratio normalized by the reference Earth ratio, shows that the depletion in volatiles is not achieved in any of the simulated planets. Most planets have a Solar photosphere ratio. On the other hand, the suitable enrichment in refractories is provided by some planets in the warm disk model.

3.5. Water delivery

In the dynamical simulations by O'Brien et al. (2006), the EJS and EEJS simulations result in a water content of Earth-analogs that is very low, since with high giant planet eccentricities the supply of planetesimals from the icy asteroid belt region are suppressed. With the same giant planet orbits, in the simulations created by Morishima et al. (2010), a larger fraction of the planets originates from the outer region of the planetesimal annulus, which should result in a overabundance of water.

Since water is one of the most important molecules in the formation and evolution of life, we focus briefly on the H abundance in the simulated planets to quantify the above statement by Morishima et al. (2010). We keep in mind that the condensation calculations will not be valid at temperatures where hydrated species form. Nevertheless, the abundance of H within a planet in our model will provide a benchmark on the abundance of material from the outer edge of the planetesimal disk.

The H abundance is mainly controlled by the condensation of serpentine. All disks are too hot for water ice to condense. Serpentine condenses out of the solar nebula in the outer region of the planetesimal annulus within 4 AU before t_{trans} depending on the disk model. This edge is located inside 4 AU only in the self-similar solution model and in the two-dimensional model. In Fig. 6, the number of planets comprising a wt.% of H is shown. In general, different disk models lead to a completely different wt.% of H since its inner edge of condensation is close to the outer edge of the dynamical simulations.

3.6. Extreme disk chemistry

In Bond et al. (2010b), the bulk chemical composition of hypothetical terrestrial planets in extrasolar planetary systems is simulated. The element abundance of stars can be significantly different compared to the abundance in our solar nebula and giant planets in other systems can have completely different masses and orbits than Jupiter and Saturn. These circumstances could lead to the formation of rather extreme objects such as carbon-rich planets.

We explore one extreme case by adopting element abundances of HD 4203 used in Bond et al. (2010b). A recent article by Fortney (2012) pointed out that the derived stellar C/O ratios could be overestimated. Nevertheless, we use the same disk models and dynamical simulations as before. If we focus on the most extreme disk models again, Fig. 7 reveals that the wt.% of C becomes extremely high for small radial distances. The main forms of carbon at these distances are C, TiC and SiC. Considering that most of the

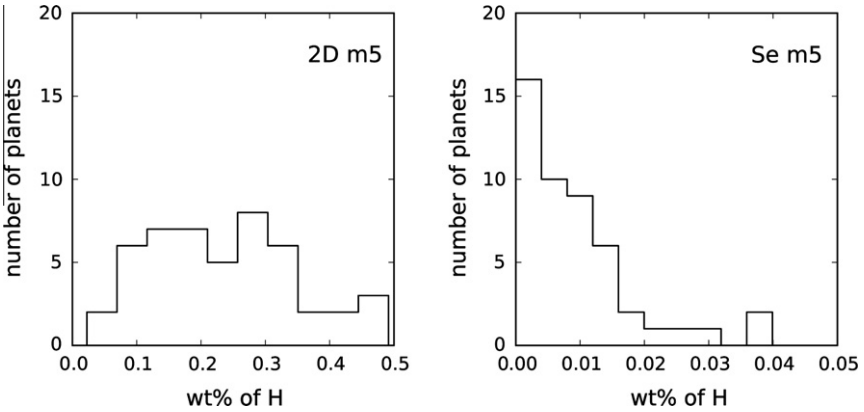

Fig. 6. The wt.% of H of all planets for the two-dimensional (2D) and the self similar solution (Se) disk models for m5 disks. All the other models give no condensation of H rich material in the disk.

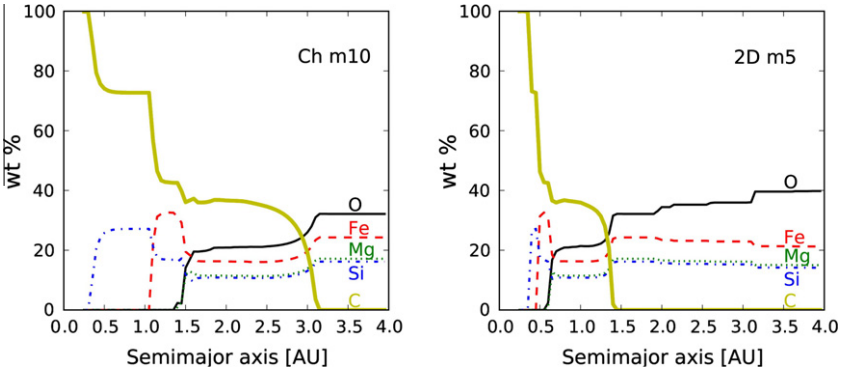

Fig. 7. These figures show the weight percentage chemical abundance of the most frequent elements in solid species as a function of radius in the case of extreme initial abundances. The models are from left to right: Chambers with m10 and two-dimensional with m5. The thick line shows a very high abundance of C close to the star. At 4 AU, the elements are from top to bottom: O, Fe, Si, Mg, C.

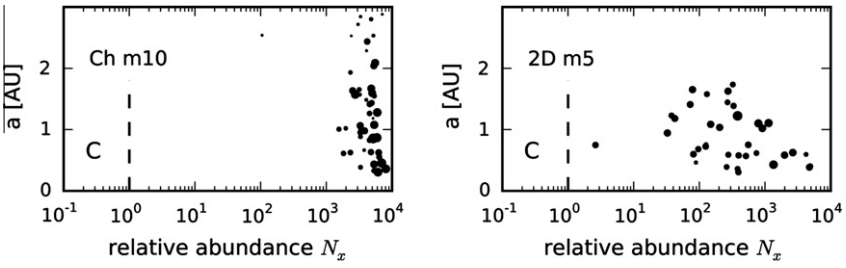

Fig. 8. The C-abundance of all planets in the case of an extreme solar nebula. Details are explained in the caption of Fig. 5 (no reference values of Mercury, Venus and Mars are shown here). Note the different scaling of the abscissa since the variations are much stronger. The left panel represents m10 with a Chambers disk model (Ch). The right panel represents m5 with a two-dimensional disk model (2D).

planetesimals are initially located inside ~ 1 AU, planetesimals that will end up in a planet do not have concentrations higher than 40 wt.%. From this, we can expect that a wt.% of 40 might be the highest bulk abundance that can be achieved in the planets. The C-abundance of the planets normalized to the Earth value is shown in Fig. 8. Note that around one quarter of all planets do not contain C and are not plotted. This also holds in the case of the two-dimensional disk model. In both disk models, there is a wide spread in the C-abundance which can be up to a factor 10^4 times higher than the Earth C-abundance. Hence, the wt.% of C in the planets range from a few to 40%. This verifies results by Bond et al. (2010b).

4. Caveats, implications and future work

In this section, we focus on the major caveats and implications that result from our approach concerning the three main goals we introduced at the beginning: First, the estimation and interpretation of abundance profiles in the circumstellar disk. Second, the interpretation of the source regions of the simulated planets in the dynamical simulations. Finally, the combination of abundance profiles and dynamical simulations, the dependence of the resulting bulk chemical compositions on the initial assumptions and an application to the Solar System.

4.1. Caveats

4.1.1. Disk abundance profiles

Strong assumptions and simplifications are required when estimating chemical abundance profiles. There are a lot of uncertainties in describing the disk physics correctly, for example the source of viscosity in the disk and the opacity. Solid migration is not taken into account in the disk models, although it is expected to occur in the actual circumstellar disks (Garaud, 2007). Here, although we used very simple models, they give rise to a broad range of temperature, pressure and surface-density profiles. The self-similar solution model we constructed is not fully self-consistent in its derivation. Since we did not use it in the further study, this caveat is not important.

All calculations performed with the HSC Chemistry code are made under the assumption of chemical equilibrium. Each solid species is treated as a separate phase, and each is considered a pure substance. Therefore, only the most stable species appear as condensates. Actually, the formation of solids can be controlled by disequilibrium condensation, at least partially. In such a case, once a solid has formed, it does not interact anymore with the system. This may occur if condensation takes place fast enough and could have a significant effect on the estimated bulk composition (Bond et al., 2010a). Other disequilibrium effects are evaporation or photo-dissociation of gaseous molecules (Ebel, 2006). At low temperatures, the assumption of chemical equilibrium is not valid anymore due to lack of diffusion in the condensed solids. Hence, we cannot make reliable predictions for the volatile content in solids and the condensation of hydrated silicates like serpentine, the only solid containing H in our calculations, might be not justified. However, condensation of some hydrated material cannot be excluded at the very outer edge of the planetesimal belt. Beside this, the radial abundance of biologically important elements like H, C and N cannot be estimated correctly with our approach. C and N rich species do not condense in any of the models. They would condense out outside of 4 AU, were the disk is cold enough. In addition, the formation of water out of H and O depends on its physical environment and the presence of H on a planet need not indicate the presence of H_2O.

The transition from the disk models to the dynamical simulations contains uncertainties. Our transition approach is based on matching the initial surface density of solids at 1 AU by adapting the time of the disk model evolution. The slope of the dynamical

simulation surface density is exactly reproduced in the self-similar model. The other models provide a flatter profile and the surface density profile deviates from the initial conditions in Morishima et al. (2010). We did not study the dependence of the abundance profiles on the transition criterion systematically. If the surface density at 1 AU is changed, as it is the case for the two different solid disk masses in the dynamical simulations, the temperature and pressure profiles and thus the element abundance in the disk change significantly. A similar effect occurs when the position of the normalization is changed. If the transition time is left as a free parameter (as in Bond et al. (2010a)), the abundance profile and thus the bulk compositions will change in a similar manner as if the planetesimal disk mass is changed. If the transition takes place later, the temperature is lower which corresponds to a smaller disk mass.

Another caveat is the gas-to-dust ratio we adopted to estimate the surface density of solids in our disk models. The Morishima et al. (2010) simulations use a gas-to-dust ratio which is up to three times higher. This high ratio would result in a very high surface density at $t_{\rm trans}$. The temperature and pressure in the disks would be very high, which moves $r_{\rm in}$ to a larger radial distance. Thus, no solids would be present in large parts of the circumstellar disk at $t_{\rm trans}$. This would be inconsistent with the initial condition of the dynamical simulations. Therefore, we claimed that a gas-to-dust ratio similar to the ISM is more reasonable. Although this results in an inconsistency with the initial conditions of the dynamical simulations, a smaller gas density in the dynamical simulations would only affect the gas-drag and type I migration. The strength of the effect of this type of migration is still a matter of debate (Schlaufman et al., 2009; Paardekooper and Papaloizou, 2009).

Both caveats concerning the transition mentioned above could be eliminated by adding a more sophisticated disk model to the dynamical simulations which could directly provide temperature and pressure profiles and such a transition would be eliminated.

4.1.2. Source regions of planets

Observations reveal that there is a very large variety in the configuration of extrasolar systems. Due to limitations in computational resources, the simulations by Morishima et al. (2010) were constructed in order to reproduce the constraints given by the Solar System and their outcomes do not represent general planetary systems. The presence of Jupiter and Saturn truncates the initial planetesimal disk at 4 AU and no icy bodies beyond this limit are taken into account. Schlaufman et al. (2009) have shown that the type-I migration rate obtained from linear theory decreases by an order of magnitude, if for example the ice line is taken into account. These recent results are not included in the dynamical simulations. In addition, the analytic expression for type-I migration used by Morishima et al. (2010) does not account for the sharp surface density transition which may occur at the inner edge of the gas disk, where the disk is truncated by e.g. magnetospheric accretion. Since these strong density gradients can act as protoplanet traps (Masset et al., 2006; Hasegawa and Pudritz, 2011), they can prevent early forming protoplanets from falling into the star. Moreover, at the pressure maximum at the inner gas disk edge, gas drag will accumulate planetesimals and the edge can also act as a planetesimal trap. These effects might have significant consequences on the source regions and the final location of the planets.

Since only one simulation per set of initial conditions was performed by Morishima et al. (2010), we are not able to do a broad study of the dependence of the source regions on the initial conditions of the dynamical simulations. The late stage of formation is a stochastic process and the source regions of two planets in one group might naturally differ significantly between each other. A much larger set of simulations is necessary to explore stochasticity and parameter space.

4.1.3. Combining chemistry and dynamics

Our transition approach is based on matching the initial surface density of solids at 1 AU by adapting the time of the disk model evolution. This should not imply that all the planetesimals in the dynamical simulations formed instantaneously at their initial positions. Yet, all planetesimals are allocated the composition of solids in the disk only at the time of transition. Actually, the formation and growth of planetesimals out of solids might significantly depend on the local density, temperature and pressure and does occur inhomogeneously along radial direction in the disk. A more sophisticated chemical model could provide condensation rates of solids, so that one could keep track of the solid fraction at every radius and integrate them over time to improve our snapshot-like approach.

The inner edge $r_{\rm in}$ of the disk of solids that condenses at $t_{\rm trans}$ lies inside the inner edge of the initial planetesimal annulus in the m10 simulations for all disk models. Thus, no solids can be allocated to these planetesimals. However, in all of the dynamical simulations, all the planetesimals initially located inside $r_{\rm in}$ are lost to the star due to gas drag and type I migration and the bulk composition of the final planets can be completely estimated in every case. Nevertheless, two caveats remain: a smaller migration rate (Paardekooper and Papaloizou, 2009; Schlaufman et al., 2009) could prevent some of the innermost planetesimals from migrating into the central star. In addition, the gravity of the planetesimals initially located inside $r_{\rm in}$ gravitational interact with their counterparts on larger radial distance. Hence, they cannot simply be ignored.

According to Albarède (2009), the inner part of the Solar System was very hot as long as gas was present and any solids that condensed were depleted in volatiles relative to the Solar photosphere. Our disk models do not predict such high temperatures at transition time and the expected volatile depletion in the case of e.g. the Earth is not achieved. However, at the inner edge of the planetesimal annulus, highly volatile depleted planetesimals form but most of them are lost into the star. As previously mentioned, modifications in the migration mechanism might result in a higher fraction of volatile depleted planetesimals that end up within planets. Following Albarède (2009), the local planetesimals were dry and the water has to come from asteroids and comets from the Jupiter–Saturn region or comets from the trans-Uranian region (Morbidelli et al., 2000). However, water–vapor absorption on silicate grains can be a mechanism to form water-rich planetesimals in the hot regions of the disk (Muralidharan et al., 2008).

The reference values for the bulk composition of the Solar System rocky planets were not reproduced in detail. Since these values are themselves based to some degree on simulations and condensation calculations, they do not necessarily provide real constraints. The bulk composition of Earth (Kargel and Lewis, 1993) is based on Best Bulk Silicate Earth (BBSE) abundance values, which are estimated using several methods that have their own uncertainties. Knowing the BBSE and the volatility trends for the elements, the core abundances and the bulk chemical composition of the Earth was estimated. Uncertainties in the BBSE values, the volatility trends and the ratio of core mass to total mass result in relative errors from 5% to 10% for most elements and up to 25% for volatiles. In Lodders and Fegley (1997), element abundances for Mars are based on meteorite samples and an oxygen isotope model. The uncertainties result in a relative error of 10%. In Morgan and Anders (1980), since the data is very limited, the compositions of Venus and Mercury are determined fully theoretically. In their method, Solar gas cools under the assumption of equilibrium condensation. The composition of a body is determined by the amount of early condensate which were most likely embodied in the final planet (Morgan et al., 1978). The uncertainties arising from this method are not quantified. In summary, the fact that we can reproduce parts of the characteristic of the chemical compositions of the

planets does not imply definitely that our model is reproducing the Solar System planets. The reference values should be seen as a guideline.

There are some further caveats resulting from the dynamical simulations: the limited number of simulations does not provide any planetary system that accurately mimics the inner Solar System. Since terrestrial planet formation is a stochastic process, a larger set of simulations could provide better candidates. The simulations adopt perfect sticking if two bodies merge, and no decrease of volatiles is taken into account. More sophisticated accretion models (Stewart and Leinhardt, 2009) propose a collision outcome depending on the collision parameters concerning energy loss and fragmentation. It ranges from a perfect accretion event to complete disruption of impactor or target. The energy release of these events is clearly high enough that volatile elements can be lost. This can be a possible explanation of the overabundance of volatile elements (Bond et al., 2010a), although there are other additional processes that can result in the depletion in volatiles in the solar nebula (Davis, 2006). Finally, we note that the rather extreme composition of Mercury, e.g. its massive core, can be explained by a collision that striped off most of Mercury's mantle and crust. No such event can take place in the Morishima et al. (2010) simulations.

4.2. Implications

4.2.1. Disk abundance profiles

Different initial conditions in the dynamical simulations result in different abundance profiles in the disk. Since the temperature and pressure in a disk is directly related to the surface density, the choice of the initial disk mass and the profile of the planetesimals in the dynamical simulations controls the distributions of condensed solids in the disk. Relatively cold disks shelter a smaller variety of solids and especially smaller abundances of volatiles like H. A massive disk of planetesimals such as the m10 simulations implies a hot disk where only a few solids will condense close to the center of the disk and no H is embodied in solids inside 4 AU. A less massive disk like the m5 disk can result in the condensation of the entire range of considered elements across the disk. The initial disk mass has an equally large or even larger effect on the resulting abundance profile than the choice of the disk model.

4.2.2. Source region of planets

Comparing the source regions of the planets in the different groups reveals that most of them do not coincide. The distributions are sometimes more peaked or more flat, but the width of the source region is always very similar. Due to migration and secular resonances, the final planets are located in the inner part of their source region, some of them are even found outside of their source region. The innermost planets tend to have source regions that are flatter. A trend is visible in the planet mass dependence of source regions: Massive planets have a flat and wide source region while small planets have a steep source region which becomes steeper at larger radial distance. This implies that massive planets consist of elements that condensed from regions that spanned a larger range of physical conditions than their less massive counterparts.

4.2.3. Combining chemistry and dynamics

The bulk chemical composition of planets produced in the dynamical simulations do not reproduce the bulk chemical composition of the rocky Solar System planets in detail. We studied two realizations of a four-planet system: In both cases, the effect of different disk models or the initial composition was not dramatic and the two dynamical simulations also gave similar results. Thus, it may be difficult to reproduce the diversity in the reference bulk chemical compositions of our Solar System in a single simulation.

In Bond et al. (2010a), the transition time is a free parameter and a different parameter space is covered than in our study. We believe that our transition criterion is more self-consistent, but both methods result in a broad range of planetary compositions. Similar to Bond et al. (2010a), we overestimated the abundance of volatile elements, implying that volatile loss cannot be ignored. Keeping in mind the differences in the dynamical simulations by Morishima et al. (2010) and O'Brien et al. (2006), we found that it is the choice of the disk model (a cold disk) that controls the amount of wet planetesimals that are accreted (rather than the orbits of the jovian planets).

A comparison of the bulk chemical composition of all simulated planets reveals the effect of the choice of the disk model. If the initial disk mass of the dynamical simulations is small (m5) and the coldest disk model is chosen (two dimensional model), most of the simulated planets share the same bulk composition. Variations are only visible in O, Na, and S for planets with small semi-major axis. In this scenario, Solar System values are almost never achieved. This means that the dynamics are not significantly affecting the abundances of this model. On the other hand, a massive disk (m10) with a hot disk model (Chambers model) provides large variations in the most refractory and volatile elements, since dynamics are important in this scenario. Again, small semi-major axis planets tend to vary more than others.

A similar trend is visible in the case of planet mass, especially for S, since more massive planets vary more in bulk composition than smaller planets. A difference was also observed for the source regions of large mass and small mass planets. Massive planets tend to be comprised of a more uniform distribution of planetesimals that formed all across the disk. Therefore they tend also to vary more in bulk composition than small mass planets, which are comprised of planetesimals from the outer disk region. A larger set of dynamical simulations is needed to quantify this difference. Note that such correlations between mass and bulk composition are not visible in the extreme abundance plot for C (Fig. 8). The C abundance dominates over a larger region than for example Al in the case of a solar nebula composition and thus, the different source regions of small and large mass planets do not play an important role.

Both the self-similar solution model and the two-dimensional disk model in case of the less massive disk, give rise to the condensation of H rich material inside the planetesimal region. Water is an important component in the formation and evolution of life as we know it. If predictions about water on habitable planets in such simulations are made, the choice of the disk model and the initial disk mass have a significant effect on the amount of H that is delivered to the planets. However, condensation calculations fail in this temperature regime and perhaps the planetesimals form dry and water has to be delivered from outside the planetesimal belt.

4.3. Future work

Given all of the uncertainties and assumptions, the main caveat in this work is the limited number of dynamical simulations that were available. A new code currently under development at the University of Zurich that will be able to perform terrestrial planet formation simulations using GPUs and a much larger number of simulations is planned to explore stochasticity and a broader parameter space. With a more sophisticated disk model in the dynamical simulations, estimating the bulk composition of planetesimals should be possible more accurately and self-consistently. A large set of dynamical simulations will also provide a good starting point for statistical studies on the source region of planets. Future studies should also include a more detailed treatment of the condensation sequence which would provide a better understanding of the evolution of solid abundances in the disk.

5. Conclusions

The combination of chemical disk models and dynamical simulations opens a new avenue for exploring planet formation models and to interpret the history of the terrestrial planets (Bond et al., 2010a). The output of the coupled disk model, its chemistry and the dynamical simulations leads directly to estimates of the bulk composition of the simulated planets. We explored the sensitivity of the radial element abundance trends in the resulting planets to the disk model assumptions, the initial conditions of the dynamical simulations and initial composition of the solar nebula. Our main conclusions are:

- Bond et al. (2010b) concluded that their estimated bulk chemical composition are in excellent agreement with observed planetary values, indicating that the models in the dynamical simulations of O'Brien et al. (2006) work properly in reproducing the bulk composition of the inner Solar System. In our study, these compositions are not reproduced in detail, resulting from the different dynamical simulations and the more restrictive but self-consistent transition from the gas disk to the planetesimals. The largest discrepancies are in the abundances of volatiles and refractories and in the bulk composition of a Mercury analog.
- The source regions of planets are unique. Large mass planets form from a broader region of the initial disk whilst lower mass planets tend to comprise of planetesimals which are initially located in the outer region of the disk.
- The element abundance profiles change significantly if different disk models or if different initial disk masses of planetesimals are used. Some elements do not condense at all if the disk model predicts a high temperature along the disk. This dependence has a significant effect on the estimated bulk chemical compositions of the planets.
- In a cold disk model, the variety in the bulk compositions is small and mixing is not important. In a hot disk model, the dynamics lead to a large variety in the bulk compositions of planets in the case of the most refractory and most volatile elements. The variation tends to be larger for massive planets and for those with small semi-major axis.

Acknowledgments

We appreciate the comments by the reviewers Franck Hersant and David Rubie. We thank Jade Bond for her inspiring work on terrestrial planet bulk compositions. We thank Pascale Garaud for a useful discussion on the disk models and James Connolly for suggestions on equilibrium calculation software. We thank Ryuji Morishima for performing the dynamical simulations on the zBox supercomputer at the University of Zurich. We thank the Planet-Z community for the fruitful discussions and for providing a platform of collaboration and we thank the University of Zurich for financial support.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.icarus.2012.09.016.

References

Agnor, C.B., Canup, R.M., Levison, H.F., 1999. On the character and consequences of large impacts in the late stage of terrestrial planet formation. Icarus 142, 219–237.
Albarède, F., 2009. Volatile accretion history of the terrestrial planets and dynamic implications. Nature 461, 1227–1233.
Andrews, S.M., Wilner, D.J., Hughes, A.M., Qi, Ch., Dullemond, C.P., 2009. Protoplanetary disk structures in Ophiuchus. Astrophys. J. 700, 1502–1523.
Armitage, P.J., 2010. Astrophysics of Planet Formation. Cambridge University Press.
Asplund, M., Grevesse, N., Sauval, A.J., 2005. The solar chemical composition. In: Barnes III, T.G., Bash, F.N. (Eds.), Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis, vol. 336. Astronomical Society of the Pacific Conference Series, pp. 25–38.
Benz, W., Anic, A., Horner, J., Whitby, J.A., 2007. The origin of Mercury. Space Sci. Rev. 132, 189–202.
Bond, J.C., Lauretta, D.S., O'Brien, D.P., 2010a. Making the Earth: Combining dynamics and chemistry in the Solar System. Icarus 205, 321–337.
Bond, J.C., O'Brien, D.P., Lauretta, D.S., 2010b. The compositional diversity of extrasolar terrestrial planets. I. In situ simulations. Astrophys. J. 715, 1050–1070.
Chamberlin, T.C., 1905. In Carnegie Institution Year Book 3 for 1904. Carnegie Inst., Washington, DC, pp. 195–234
Chambers, J.E., Wetherill, G.W., 1998. Making the terrestrial planets: N-body integrations of planetary embryos in three dimensions. Icarus 136, 304–327.
Chambers, J.E., 2009. An analytical model for the evolution of a viscous, irradiated disk. Astrophys. J. 705, 1206–1214.
Chiang, E., Youdin, A.N., 2010. Forming planetesimals in solar and extrasolar nebulae. Annu. Rev. Earth Planet. Sci. 38, 493–522.
D'Alessio, P., Canto, J., Nuria, C., Lizano, S., 1998. Accretion disks around young objects. I. The detailed vertical structure. Astrophys. J. 500, 411–427.
Davis, A.M., 2006. Volatile evolution and loss. In: Lauretta, D.S., McSween, H.Y., Jr. (Eds.), Meteorites and the Early Solar System II. University of Arizona press, Tucson, AZ, USA, pp. 295–307.
Drouart, A., Dubrulle, B., Gautier, D., Robert, F., 1999. Structure and transport in the solar nebula from constraints on deuterium enrichment and giant planets formation. Icarus 140, 129–155.
Dullemond, C.P., Hollenbach, D., Kamp, I., D'Alessio, P., 2006. Models of the structure and evolution of protoplanetary disks. In: Reipurth, V.B., Jewitt, D., Keil, K. (Eds.), Protostars and Planets, vol. 951. University of Arizona Press, Tucson, pp. 555–572.
Ebel, D.S., 2006. Condensation of rocky material in astrophysical environments. In: Lauretta, D.S., McSween, H.Y., Jr. (Eds.), Meteorites and the Early Solar System II. University of Arizona press, Tucson, AZ, USA, pp. 253–277.
Fortney, J.J., 2012. On the carbon-to-oxygen ratio measurement in nearby Sun-like stars: Implications for planet formation and the determination of stellar abundances. Astrophys. J. (Letters). 747, L27.
Garaud, P., 2007. Growth and migration of solids in evolving protostellar disks. I. Methods and analytical tests. Astrophys. J. 671, 2091–2114.
Hasegawa, Y., Pudritz, R.E., 2011. The origin of planetary system architectures: I. Multiple planet traps in gaseous disc. Mon. Not. R. Astron. Soc. 417, 1236–1259.
Hartmann, L., 1998. Accretion Processes in Star Formation. Cambridge University Press.
Hayashi, C., 1981. Structure of the solar nebula, growth and decay of magnetic fields and effects of magnetic and Turbulent viscosities on the nebula. Progress of Theoretical Physics Supplement. 70, 35–53.
Hersant, F., Gautier, D., Huré, 2001. A two-dimensional model for the primordial nebula constrained by D/H measurements in the Solar System: Implications for the formation of giant planets. Astrophys. J. 554, 391–407.
Huré, J.-M., 2000. On the transition to self-gravity in low mass AGN and YSO accretion discs. Astron. Astrophys. 358, 378–394.
Huré, J.-M., Galliano, F., 2001. The global structure of thin, stratified alpha-discs and the reliability of the one layer approximation. Astron. Astrophys. 366, 359–362.
Johansen, A., Oishi, J.S., Mac Low, M.-M., Klahr, H., Henning, T., Youdin, A., 2007. Rapid planetesimal formation in turbulent circumstellar disks. Nature 448, 1022–1025.
Latimer, W.M., 1950. Astrochemical problems in the formation of the Earth. Science 112, 101–104.
Lodders, K., 2003. Solar system abundances and condensation temperatures of the elements. Astrophys. J. 591, 1220–1247.
Lodders, K., Fegley, B., 1997. An oxygen isotope model for the composition of Mars. Icarus 126, 373–394.
Lissauer, J.J., 1993. Planet formation. Ann. Rev. Astron. Astrophys. 31, 129–174.
Lynden-Bell, D., Pringle, J.E., 1974. The evolution of viscous discs and the origin of the nebular variables. Mon. Not. R. Astron. Soc. 168, 603–637.
Kargel, J.S., Lewis, J.S., 1993. The composition and early evolution of Earth. Icarus 105, 1–25.
Kitamura, Y., Momose, M., Yokogawa, S., Kawabe, R., Tamura, M., Ida, S., 2009. Investigation of the physical properties of protoplanetary disks around T Tauri stars by a 1 arcsecond imaging survey: Evolution and diversity of the disks in their accretion stage. Astrophys. J. 581, 357–380.
Kokubo, E., Kominami, J., Ida, S., 2006. Formation of terrestrial planets from protoplanets. I. Statistics of basic dynamical properties. Astrophys. J. 642, 1131–1139.
Masset, F.S., Morbidelli, A., Crida, A., Ferreira, J., 2006. Disk surface density transitions as protoplanet traps. Astrophys. J. 642, 478–487.
Morgan, J.W., Anders, E., 1980. Chemical composition of Earth, Venus and Mercury. Proc. Natl. Acad. Sci. 77, 6973–6977.
Morgan, J.W., Hertogen, J., Anders, A., 1978. The Moon – Composition determined by nebular processes. Moon Planets 18, 465–478.

Morbidelli, A. et al., 2000. Source regions and time scales for the delivery of water to Earth. Meteorit. Planet. Sci. 35, 1309–1320.
Morishima, R., Stadel, J., Moore, B., 2010. From planetesimals to terrestrial planets: N-body simulations including the effects of nebular gas and giant planets. Icarus 207, 517–535.
Muralidharan, K., Deymier, P., Stimpfl, M., de Leeuw, N.H., Drake, M.J., 2008. Origin of water in the inner Solar System: A kinetic Monte Carlo study of water adsorption on forsterite. Icarus 198, 400–407.
O'Brien, D.P., Morbidelli, A., Levison, H.F., 2006. Terrestrial planet formation with strong dynamical friction. Icarus 184, 39–58.
Palme, H., O'Neill, H., St, C., 2003. Cosmochemical estimates of mantle composition. Treatise on geochemistry. In: Carlson, R.W. (Ed.), The Mantle and Core, vol. 2. Elsevier-Pergamon, Oxford, pp. 1–38.
Pasek, M.A., Milsom, J.A., Ciesla, F.J., Lauretta, D.S., Sharp, C.M., Lunine, J.I., 2005. Sulfur chemistry with time-varying oxygen abundance during Solar System formation. Icarus 175, 1–14.
Paardekooper, S.-J., Papaloizou, J.C.B., 2009. On corotation torques, horseshoe drag and the possibility of sustained stalled or outward protoplanetary migration. Mon. Not. R. Astron. Soc. 394, 2283–2296.
Papaloizou, J.C.B., Terquem, C., Nelson, R.P., 1999. Disc and planetary formation. In: Sellwood, J.A., Goodman, J. (Eds.), Astrophysical Discs, vol. 160. Astronomical Society of the Pacific Conference Series, pp. 186–206.
Raymond, S.N., Quinn, Th., Lunine, J.I., 2004. Making other earths: Dynamical simulations of terrestrial planet formation and water delivery. Icarus 168, 1–17.
Raymond, S.N., O'Brien, D.P., Morbidelli, A., Kaib, N.A., 2009. Building the terrestrial planets: constrained accretion in the inner Solar System. Icarus 203, 644–662.
Righter, K., Drake, M.J., Scott, E.R.D., 2006. Compositional relationships between meteorites and terrestrial. In: Lauretta, D.S., McSween, H.Y., Jr. (Eds.), Meteorites and the Early Solar System II. University of Arizona press, Tucson, AZ, USA, pp. 803–828.
Rubie, D.C. et al., 2011. Heterogeneous accretion, composition and core–mantle differentiation of the Earth. Earth Planet. Sci. Lett. 301, 31–42.
Ruden, S.P., Pollack, J.B., 1991. The dynamical evolution of the protosolar nebula. Astrophys. J. 375, 740–760.
Safronov, V.S., 1969. Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets. Moskau, Nauka. Engl. Transl. NASA TTF-677, 1972.
Schlaufman, K.C., Lin, D.N.C., Ida, S., 2009. The signature of the ice line and modest type I migration in the observed exoplanet mass-semimajor axis distribution. Astrophys. J. 691, 1327–1332.
Shakura, N.I., Sunyaev, R.A., 1973. Black holes in binary systems. Observational appearance. Astron. Astrophys. 24, 337–355.
Stepinski, T.F., 1998. The solar nebula as a process – An analytic model. Icarus 132, 100–112.
Stewart, S.T., Leinhardt, Z.M., 2009. Velocity-dependent catastrophic disruption criteria for planetesimals. Astrophys. J. 691, 133–137.
Touboul, M., Kleine, T., Bourdon, B., Palme, H., Wieler, R., 2007. Late formation and prolonged differentiation of the Moon inferred from W isotopes in lunar metals. Nature 450, 1206–1209.
Tsiganis, K., Gomes, R., Morbidelli, R., Levision, H.F., 2005. Origin of the orbital architecture of the giant planets of the Solar System. Nature 435, 459–461.
Walsh, K.J., Morbidelli, A., Raymond, S.N., O'Brien, D.P., Mandell, A.M., 2011. A low mass for Mars from Jupiter's early gas-driven migration. Nature 475, 206–209.
Weidenschilling, S.J., 1980. Dust to planetesimals – Settling and coagulation in the solar nebula. Icarus 44, 172–189.
White, W.B., Johnson, S.M., Dantzig, G.B., 1958. Chemical equilibrium in complex mixtures. J. Chem. Phys. 28, 751–755.

5

Paper III
Stability of Orbits

In this final paper, we predicted orbits of hypothetical habitable Super-Earths in known multi planet systems by means of dynamical stability studies. This paper was motivated mainly by two circumstances. First, the number of observed planetary systems increases continuously. These days, the first potentially habitable planets are discovered, and there will be more in the next years. This search can be guided by predictions based on dynamical simulations. Second, the newly developed GPU code GENGA [101] provides the possibility to run more than thousand N-body simulations on one GPU in parallel. This allows to carry out numerous simulations of the long-term dynamical evolution of planetary systems. Hence, this study provides a good first application to profit from the advantages of this new code.

We found that most of the systems we took into account harbor potentially a Super-Earth or Earth-like planet in the habitable zone. Upcoming observations should confirm our predictions.

This paper was published 2013 in Monthly Notices of the Royal Astronomical Society MNRAS [118].

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS Advance Access published June 14, 2013

MNRAS (2013)

doi:10.1093/mnras/stt883

Super-Earths and dynamical stability of planetary systems: first parallel GPU simulations using GENGA

S. Elser,* S. L. Grimm and J. G. Stadel

Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Accepted 2013 May 16. Received 2013 May 16; in original form 2013 April 2

ABSTRACT

We report on the stability of hypothetical super-Earths in the habitable zone of known multi-planetary systems. Most of them have not yet been studied in detail concerning the existence of additional low-mass planets. The new *N*-body code GENGA developed at the University of Zürich allows us to perform numerous *N*-body simulations in parallel on graphics processing units. With this numerical tool, we can study the stability of orbits of hypothetical planets in the semimajor axis and eccentricity parameter space in high resolution. Massless test particle simulations give good predictions on the extension of the stable region and show that HIP 14810 and HD 37124 do not provide stable orbits in the habitable zone. Based on these simulations, we carry out simulations of 10 $M_\oplus$ planets in several systems (HD 11964, HD 47186, HD 147018, HD 163607, HD 168443, HD 187123, HD 190360, HD 217107 and HIP 57274). They provide more exact information about orbits at the location of mean motion resonances and at the edges of the stability zones. Beside the stability of orbits, we study the secular evolution of the planets to constrain probable locations of hypothetical planets. Assuming that planetary systems are in general closely packed, we find that apart from HD 168443, all of the systems can harbour 10 $M_\oplus$ planets in the habitable zone.

Key words: methods: numerical – celestial mechanics – planets and satellites: dynamical evolution and stability.

1 INTRODUCTION

In the past two decades, numerous planetary systems have been discovered (Schneider et al. 2011). Most of those systems contain only a single known planet. Since Butler et al. (1999) announced the discovery of the first multiple planet system around a normal star, many multiple planetary systems were discovered and confirmed (Wright 2010). Many more will follow in the next few years when a high percentage of the present Kepler candidate planets are going to be confirmed (Borucki et al. 2011). There are planetary systems with up to six planet candidates (Lissauer et al. 2011; Tuomi et al. 2013). Both the Doppler spectroscopy and the detection via transit observations prefer massive, respectively, large planets close to the host star. The discovery of Earth-like planet candidates with respect to mass and size has just started thanks to the high precision spectrograph HARPS (Pepe et al. 2011; Dumusque et al. 2012) or space missions like Kepler (Borucki et al. 2012; Fressin et al. 2012), whereas planets of several Earth-masses, so called super-Earths, were discovered in the habitable zone (HZ) of stars (Vogt, Butler & Haghighipour 2012; Lo Curto et al. 2013). Nevertheless, the detection of an Earth-like planets in the HZ around a Sun-like star is extremely difficult and was not yet successful.

To guide the search for additional planets in known planetary systems, numerical stability studies are a powerful tool. In the recent years, numerical investigations estimated stability zones in known systems which might harbour unknown planets (Menou & Tabachnik 2003; Asghari et al. 2004; Barnes & Raymond 2004; Raymond & Barnes 2005; Hinse et al. 2008; Kopparapu et al. 2008; Fang & Margot 2012). Barnes & Raymond (2004) and Raymond & Barnes (2005) had shown the location of a stable zone in the 55 Cancri system before planet *f* was discovered right at the inner edge of this zone (e.g. Fischer et al. 2008). They also predicted the existence of a Saturn-mass planet in HD 74156, which was later discovered by Bean et al. (2008). However, this prediction of the orbit and mass of an extra planet is not yet confirmed and under debate (Baluev 2009; Wittenmyer et al. 2009).

Many multiple planetary systems tend to be near the edge of stability and small perturbations would destabilize the system (e.g. Barnes & Quinn 2004). The 'packed planetary systems' (PPS) hypothesis (Barnes & Raymond 2004) claims that every stable region between two neighbouring (known) planets is occupied by an additional (unknown) planet. Hence, all planetary systems tend to form 'dynamically full' and have no large gaps between the planets. Based on this hypothesis, stability regions that are identified

*E-mail: selser@physik.uzh.ch

in between known planets should potentially host additional planets. Most likely, those planets are not very massive and the impact of their perturbation on the known planet orbits might be smaller than the observational limit. Hence, they cannot be deduced from residuals in current (Doppler spectroscopy) data.

There is a major drawback when studying the stability regions in present day planetary system configurations, because we do not take into account the effects of potential early evolution of the known planets on the formation and evolution of the hypothetical planets. Despite this, early migration of giant planets through the initial planetesimal belt need not inhibit the formation of terrestrial planets, as long as the migration time-scale is small (e.g. Mandell & Sigurdsson 2003; Raymond, Mandell & Sigurdsson 2006). Dynamical instability of the initial giant planet configuration may result in ejection of one of the giants or in a merger with the central star. Such events might strongly affect the stability of a hypothetical super-Earth-sized planet located in the stability region of the final giant planet configuration and would also explain the high eccentricities of many of the observed planets. Hence, the width of stable regions in the parameter space are overestimated when dynamical instability played a significant part in forming the final giant planet configuration (Matsumura, Ida & Nagasawa 2013). However, if we assume that some hypothetical planets might form or survive despite of the early evolution of the known planets in a system, they would be found in the stability zones studied in this work.

The goal of this study is the prediction of stable orbits in the HZ of various extra solar multiple planetary systems, most of which have not yet been studied in much detail concerning the stability of hypothetical planets. As a major selection criterion, we chose systems whose inner- and outermost observed planets (partially) enclose the HZ of the system. To calculate the orbital movement of the planets, we use a new code developed at the University of Zürich called GENGA (Grimm & Stadel, in preparation), which runs completely on a graphics processing unit (GPU). This simulation code allows either a single integration with many bodies (up to ten thousand massive bodies and hundreds of thousands of massless test particles), or many parallel integrations of systems with fewer bodies to be performed on a GPU. We start to constrain stable regions in the parameter space of semimajor axis and eccentricity of a hypothetical planet analytically based on the present planets orbits. This is the first indicator on the presence of a stable zone in the initial parameter space. Then, we integrate the orbits of massless test particles in the HZ of the planetary systems. Finally, we focus on the identified stability regions and perform a large number of simulations to explore the parameter space in more detail. In this case, each simulation contains the known planets plus a massive hypothetical test planet. The stability of the test planet and its perturbations on the known planets indicate if a massive planet can be present in the HZ. All told, these simulations required around 2500 GPU-days or two months of wall-clock time on our CPU cluster.

This work is structured as follows: in Section 2, we present the systems that we take into account. Moreover, analytic approaches to estimate the stability of a planetary system are briefly presented. Then, in Section 3, we introduce GENGA and show some comparisons with similar codes to highlight the advantages of this powerful new tool. In addition, we present the setup for the simulations with massless particles and massive hypothetical super-Earths. Section 4 shows the main results. Besides presenting the extent of the stability region in each system, we highlight the most important insights and constrain the most likely regions where hypothetical super-Earths may still be found. Finally, we conclude this work in Section 5.

2 DATA AND METHODS

First, our data sample is described and we explain our motivation to choose this set of systems. Then, the PPS hypothesis is briefly described. Analytic methods to predict stable orbit locations in the semimajor axis and eccentricity parameter space are shown.

2.1 Data sample selection

The search for habitable planets is one of the main goals of present day astronomy. A habitable planet is often described as a terrestrial planet of the order of the mass of the Earth up to the mass of a super-Earth ($\approx 10\,M_\oplus$) located in the HZ of its host star. The HZ of a star is given by an annulus in distance where water on the surface of a planet can sustain in liquid form. A more general concept that takes into account the average time of a planet spending in the HZ is the eccentric habitable zone (EHZ). The exact definitions that are used in this work are given in appendix A.

We focus on systems in which the HZ is enclosed between the orbits of the innermost and outermost planets. If the HZ is enclosed only partially, the enclosed fraction should be significant, that means more than half of the HZ. Otherwise, most planets initially located inside the HZ will be perturbed or crash with the known planet. Focusing on such systems with (partially) enclosed HZ, the parameter space of interest is limited by the planets in the system and its HZ. If the PPS-hypothesis holds, every stable zone we find should potentially harbour (at least) an additional planet as a consequence of the systems formation process. The sample we use is shown in Table 1. Our sample does not represent all known multiplanetary systems with an (partially) enclosed HZ. In order to produce new results and save computational resources, we focus on systems that have not yet been studied in detail concerning stable region in the HZ (beside HD 47186, which allows a direct comparison of our simulation method). Hence, we exclude systems like 55 Cancri or HD 74156, which would also correspond to our selection criterion. In addition, we do not take into account any Kepler candidate systems.

2.2 Analytic predictions

Before studying the planetary systems with numerical methods, we present some analytic approaches with various levels of complexity to constrain and to quantify the stability in a system. Although none of them can predict details on the stability region in the (a,e)-plane, they are by far less time consuming and are the first step when studying a system.

In the case of two-planet systems, an analytic stability boundary (Barnes & Greenberg 2006, 2007) can be calculated, which is based on fundamental quantities of the system. Following Marchal & Bozis (1982) and Gladman (1993), the system is called Hill-stable and the orbits of the planets will never cross, if the ratio $\beta/\beta_{\rm crit}$ is larger than unity. β is a quantity that depends on the energy and the total angular momentum of the system, $\beta_{\rm crit}$ depends only on the masses of the star and planets:

$$\beta = \frac{-2(M_* + M_1 + M_2)}{G^2(M_1M_2 + M_*M_1 + M_*M_2)^3} L^2 E \tag{1}$$

$$\beta_{\rm crit} = 1 + \frac{3^{4/3} M_1 M_2}{M_*^{2/3}(M_1 + M_2)^{4/3}} - \frac{M_1M_2(11M_1 + 7M_2)}{3M_*(M_1 + M_2)^2} + \cdots, \tag{2}$$

where M_*, M_1 and M_2 are the masses of the star and two planets, given that $M_1 > M_2$. Here G is the gravitational constant and E

Table 1. Planetary systems of this study. The stellar mass (M_*), the stellar surface temperature (T_*) and the stellar radius (R_*) are shown. For each planet in the system, the minimum mass ($m\sin i$), the semimajor axis (a) and the eccentricity (e) are listed. Data from exoplanets.org (Wright et al. 2011) in 2012 September 18.

Star	M_* ($M_\odot$)	T_* (K)	R_* ($R_\odot$)	Planet	$m\sin i$ ($M_{\rm jup}$)	a (au)	e
HIP 14810	0.99	5485	1.32	b	3.874	0.0692 ± 0.001 15	0.142 48 ± 0.000 95
				c	1.275	0.5454 ± 0.0091	0.153 ± 0.0132
				d	0.581	1.886 ± 0.036	0.165 ± 0.04
HD 37124	0.85	5500	0.77	b	0.674	0.5336 ± 0.0089	0.054 ± 0.028
				c	0.648	1.710 ± 0.029	0.125 ± 0.055
				d	0.687	2.807 ± 0.06	0.16 ± 0.14
HD 163607	1.09	5543	1.7	b	0.769	0.3592 ± 0.006	0.730 ± 0.02
				c	2.292	2.418 ± 0.041	0.120 ± 0.06
HIP 57274	0.73	4640	0.68	b	0.037	0.0713 ± 0.00163	0.19 ± 0.1
				c	0.41	0.1778 ± 0.0041	0.050 ± 0.02
				d	0.529	1.007 ± 0.027	0.270 ± 0.05
HD 190360	0.983	5552	1.08	b	1.535	3.973 ± 0.071	0.313 ± 0.0191
				c	0.059	0.1292 ± 0.0022	0.237 ± 0.082
HD 147018	0.927	5441	1.053	b	2.127	0.2389 ± 0.004	0.4686 ± 0.0081
				c	6.593	1.923 ± 0.039	0.133 ± 0.011
HD 168443	0.995	5491	1.59	b	7.697	0.2938 ± 0.0049	0.529 ± 0.024
				c	17.386	2.853 ± 0.048	0.2113 ± 0.00171
HD 11964	1.107	5349	1.67	b	0.618	3.155 ± 0.059	0.041 + 0.088/−0
				c	0.078	0.2285 ± 0.0038	0.30 ± 0.17
HD 47186	0.99	5675	1.13	b	0.071	0.049 84 ± 0.000 83	0.038 ± 0.02
				c	0.348	2.387 ± 0.078	0.249 ± 0.073
HD 217107	1.108	5704	1.5	b	1.401	0.0750 ± 0.001 25	0.1267 ± 0.0052
				c	2.615	5.33 ± 0.2	0.517 ± 0.033
HD 187123	1.037	5815	1.14	b	0.51	0.042 09 ± 0.0007	0.0103 ± 0.0059
				c	1.942	4.83 ± 0.37	0.252 ± 0.033

Table 2. Values of $\beta/\beta_{\rm crit}$ of planetary systems in this studies. Systems with more than two known planets are marked with (a). In this case, the planet pair enclosing the HZ is taken into account.

Star	Pair	$\beta/\beta_{\rm crit}$
HIP 14810[a]	c–d	1.245
HD 37124[a]	b–c	1.248
HD 163607	b–c	1.575
HIP 57274[a]	c–d	1.581
HD 190360	b–c	1.781
HD 147018	b–c	1.806
HD 168443	b–c	2.005
HD 11964	b–c	2.041
HD 47186	b–c	6.134
HD 217107	b–c	8.941
HD 187123	b–c	15.091

and L are the total energy and orbital angular momentum of the system. This ratio, shown in Table 2 for each system, can be used to predict the possible existence of additional planets. According to Barnes & Greenberg (2007), numerical simulations have shown that $\beta/\beta_{\rm crit} \lesssim 1.5$ indicates that the system tends to be fully packed, whereas a system with $\beta/\beta_{\rm crit} \gtrsim 2$ offers stable zones for additional unknown planets. For $1.5 \lesssim \beta/\beta_{\rm crit} \lesssim 2.0$, it is not clear if the system is packed or not. The four systems that contain more than two known planets are also listed. It is not guaranteed that $\beta/\beta_{\rm crit} = 1$ means Hill stability of any individual pair. The above limits hold if the additional planets in the system are small (e.g. HIP 57274) or well separated compared to the pair that is taken into account for calculating $\beta/\beta_{\rm crit}$. Based on the above argument, we expect stable orbits in all systems apart from HIP 14810 and HD 37124.

The main osculating elements of the known planets constrain the osculating elements of any hypothetical planet. A test particle whose initial orbit crosses that of a planet is highly in danger of colliding or being scattered out of the system as a result of a close encounter. The location of crossing orbits are given by the point in the (a,e)-plane where pericentre (resp. apocentre) and apocentre (resp. pericentre) of a particle and a planet coincide. These limits provide a very general constraint on the size and shape of the stability region in the (a,e)-plane.

Capture in low-order mean motion resonance (MMRs) can provide stability beyond the crossing orbit of the planets, (e.g. Kopparapu et al. 2008; Raymond, Barnes & Gorelick 2008), or destabilize planets in the stability region. More important, the zone of the dynamical influence of a planet is larger than its physical cross-section. This gravitational zone of influence of a planet i is often expressed as some factor c_i times the Hill radius (Hamilton & Burns 1992),

$$R_{{\rm Hill},i} \equiv \left(\frac{M_{p,i}}{3M_*}\right)^{1/3} a_i, \tag{3}$$

where $i = 1, 2$ refers to the enclosing planets. Without loss of generality, we take into account a two-planet system where (a_1, e_1) are the osculating elements of the inner planet and (a_2, e_2) the corresponding elements of the outer planet. The lines of crossing orbits in (a,e)-space are given by

$$a_1(1+e_1) + c_1 R_{{\rm Hill},1} = a(1-e), \tag{4}$$

$$a(1+e) = a_2(1-e_2) - c_2 R_{{\rm Hill},2}. \tag{5}$$

In general, the factors of c_i are unknown and $c_1 = c_2 = 0$ provides a first insight. To account for the dynamical influence, a common choice is $c_1 = c_2 = 3$ (Menou & Tabachnik 2003) or higher.Studying

Kepler systems with two known planets, Fang & Margot (2012) obtained $c_1 = 19.4$ for accounting the influence of the inner planet outwards and $c_2 = 4.2$ for accounting the influence of the outer planet inwards. Jones, Sleep & Underwood (2006) used cubic fits on c_1 and c_2 obtained from simulations to get the factors for any planetary system by interpolation. They found $1 \lesssim c_2 \lesssim 3$, decreasing with planet eccentricity and $3 \lesssim c_1 \lesssim 13$, increasing with planet eccentricity. For our purpose, c_1 and c_2 are estimated by solving the system of equations (4) and (5) for e to get a piecewise function $e = e(a, c_1, c_2)$. Then, this function is fitted to the edge of the stable regions in the (a,e)-plane, which gives c_1 and c_2. The maximum eccentricity of all stable particles, $e_{\rm top}$, is then given by the maximum of the function $e = e(a, c_1, c_2)$ and can be interpreted as a measure of the stable zone. We use this to estimate $e_{\rm top}$ and to check how well the edges of the stability region can be expressed by equations (4) and (5).

Since analytic estimates are limited and e.g. the estimation of the correct c_i depends on numerical studies, we directly focus on N-body simulations to find stable orbits.

3 SIMULATIONS

Similar to Raymond et al. (2008) we use the term 'test planets' for massive bodies which fully interact with the planets in contrast to the massless 'test particles' which trace only the gravitational potential of the planets. To test the stability of planets, in a first attempt we used massless test particles which are computationally less expensive than simulations with massive test planets.

The main orbital elements of the current best-fitting orbits of the known planets are given in Table 1. The minimum mass, semimajor axis a and eccentricity e are shown with their observational uncertainties, which we use in the further study of the results. We randomly choose a mutual inclination of $i < 1^\circ$ and a longitude of the ascending node randomly distributed from 0° to 360°. The argument of periastron and the time of periastron passage are given by the references. If the actual inclination of the system were larger than a few degrees, the planet masses would be much larger and would change the dynamics of the system significantly.

In every planetary system, the two planets enclosing the HZ are named as follows: the planet whose semimajor axis is smaller than the centre of the HZ (A6) is called the 'inner' planet, the planet whose semimajor axis is larger than the centre of the HZ is called 'outer' planet.

In each simulation, the goal is to conserve energy up to 1 part in 10^5. This is achieved by choosing a suitable combination of time step and order of the integrator for each system, Table 3. The maximum time step is preset to 2 d.

3.1 The GPU Code GENGA

Modern graphics cards and the specialized variants for pure computing found in supercomputers such as the Cray XK series can perform a large number of operations in parallel by launching a large number of execution threads. The limitation is that these threads are not independent and should perform, as much as possible, the same instruction on different data (SIMD). This type of high performance computing based on GPUs can speed up many numerical tasks by a large factor over what is possible on a CPU as long as enough parallel work is available. The simulation of planetary systems would seem to provide enough parallelism as long as enough bodies are involved in the simulation ($\gtrsim$100) or enough independent systems are evolved simultaneously. Since the memory transfer between the CPU and GPU is currently still a bottleneck, GENGA runs completely on the GPU where it can take advantage of the very fast, but limited, memory that exists there. Only the outputs are transferred back to the CPU. GENGA is implemented in CUDA C by Grimm and Stadel and runs on NVIDIA GPUs with compute capability 2.0 or higher. A detailed paper is in preparation, but we will briefly present a few aspects of this new code in the following.

The GENGA Code is a hybrid symplectic integrator, based on the MERCURY code (Chambers 1999). The symplectic integrator is a mixed variable integrator as described by Wisdom & Holman (1991); Saha & Tremaine (1992). It integrates the planetary orbits for a large time-scale with a very good energy conservation. Gravitational interactions between planets are computed as perturbations of the Keplerian orbits. If two planets are in a close encounter, the perturbation potential becomes dominant and the integrator breaks down. The hybrid symplectic integrator switches in these cases smoothly to a direct N-body Bulirsch–Stoer integrator which integrates the close encounter phase up to machine precision. Two planets are in a close encounter when their separation r_{ij} is less than a critical radius, defined as

$$r_{\rm crit} = \max(r_{{\rm crit},i}, r_{{\rm crit},j}), \tag{6}$$

with

$$r_{{\rm crit},i} = \max(3R_{{\rm Hill},i}, 0.4\tau v_{\rm max}), \tag{7}$$

where τ is the time step. In the GENGA code we generalized the second-order symplectic integrator to fourth and sixth order, as described by Yoshida (1991). The higher orders are especially a

Table 3. The massive test planet simulations in detail. The time step Δt and the order of the integrator $\mathcal{O}$ are two parameters that control the accuracy of the simulation. The number of $N_{\rm init}$ simulations are sampled equally spaced in the (a,e)-plane in $a_{\rm min} \le a \le a_{\rm max}$ and $0 \le e \le e_{\rm max}$. $N_{\rm stab}$ is the number of test planets that are on a stable orbit. $N_{\rm Hill}$ is the number of planets on a stable orbit that experience a close encounter. $f_{\rm stab}$ is the percentage of stable simulations in the massive test planet simulations. $F_{\rm stab}$ normalizes $f_{\rm stab}$ to the area between the planets and $0 < e < 1$. $F_{\rm stab,m=0}$ is the normalized percentage in the massless test particle simulations.

System	Δt (d)	$\mathcal{O}$	$a_{\rm min}$	$a_{\rm max}$	$e_{\rm max}$	$N_{\rm init}$	$N_{\rm stab}$	$N_{\rm Hill}$	$f_{\rm stab}$ (per cent)	$F_{\rm stab}$ (per cent)	$F_{\rm stab,m=0}$ (per cent)
HD 163607	0.80	4	0.50	2.0	0.5	5000	680	24	13.6	4.9	5.5
HD 217107	0.35	4	0.10	3.0	0.8	5000	1883	0	39.5	17.3	16.4
HIP 57274	0.40	2	0.20	0.7	0.5	5000	1149	1	23.2	6.2	5.7
HD 11964	2.00	4	0.25	2.8	0.7	5000	2986	129	61.4	37.4	34.8
HD 187123	0.50	2	0.20	3.5	0.8	5000	2891	30	56.0	34.0	32.6
HD 147018	1.00	4	0.30	1.3	0.5	5000	729	113	15.0	4.4	5.3
HD 47186	0.50	4	0.06	2.3	0.65	5000	2205	122	55.9	35.1	32.4
HD 168443	0.30	2	0.70	1.5	0.4	5000	49	48	0.9	0.1	1.1
HD 190360	0.85	2	0.15	2.5	0.8	5000	2409	11	48.2	23.5	23.0

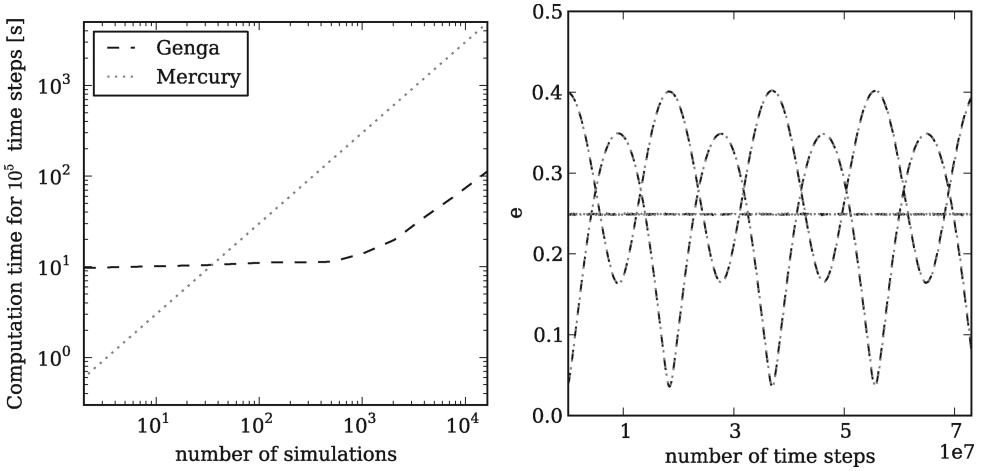

Figure 1. Performance of the GENGA Code. Left-hand panel: comparison of the performance between GENGA on one GPU (dashed line) and MERCURY on one CPU (dotted) line. Right-hand panel: comparison of a simulation output. The secular evolution of the eccentricities of a three-planet system is shown. It is HD 47186 with a 10 $M_\oplus$ test planet initially located at $(a, e) = (0.2\,\mathrm{au}, 0.4)$. MERCURY and GENGA are in nearly perfect agreement here.

good choice if the innermost planet has a very small semimajor axis and a high eccentricity.

We use the GENGA Code in two different modes: First, to simulate the planetary systems with a large number of test particles, and second, to simulate a large number of small, independent, planetary systems with different configurations. In the test particle mode we use one CUDA thread per test particle, in the multi simulation mode we use one CUDA thread per body. Fig. 1 shows the computation time for GENGA (on a NVIDIA Geforce GTX 590 graphic card) and MERCURY (on an Intel Xeon 2.8 GHz CPU) to simulate a set of three body simulations. At a low number of simulations, the GPU overhead dominates and MERCURY is faster. At a high number of simulations, GENGA benefits from the large number of GPU cores. At around 1000 simulations, the GPU is fully occupied, and the computation time begins to increase. At 16 000 simulations the GPU is about 40 times faster than one CPU.

The massive test planet simulations of each system are split on to four GPUs in most cases. This results in 1250 simulations per GPU, which allows the maximum efficiency of the code. The computation time depends on different factors: integration time step and order of the symplectic integrator, mainly controlled by the innermost planet, the survival rate of test planets and the number of close encounters. The minimum wall-clock time is around 120 d for HD 147018 and the maximum wall-clock time is around 800 d for HD 47186.

3.2 Massless test particle simulations

In the test particle simulation, we placed 20 000 test particles equally spaced in 500 steps between the semimajor axis of the inner planet a_{inner} and the semimajor axis of the outer planet a_{outer} and equally spaced in 40 steps in $0.0 \leq e \leq 0.8$. The inclinations are assigned randomly under the condition $i < 1^\circ$. The argument of periastron, the longitude of the ascending node and the mean anomaly are drawn randomly between 0° and 360°.

A test particle is representing an unstable orbit if it collides with one of the known planets or if its distance to the star exceeds 20 au. We stop the simulations when the overall shape of the orbital zone is visible, this means when the rate of orbits becoming unstable decreases significantly. This takes place after a few Myr and gives a rough idea of the stable zone and its features. Hence, we define stability by the survival of a planet.

3.3 Massive planets

The initial sampling of the (a,e)-space in the case of massive test planets is guided by the results in the massless test particle simulations. The minimum and maximum a and the maximum e of the surviving test particles are approximately taken as limits. The extent of the sampled regions and further simulation details are given in Table 3. We run each simulation for 10 Myr. Most of the unstable orbits will be identified in 10^6 orbits (Barnes & Raymond 2004). The conditions for an orbit to be identified as unstable are the same as in the test particle simulations discussed previously.

Simulations with massive test planets provide additional information about the stability of planets in the system. Depending on the mass of the test planet, the orbital parameters of the other planets (and the star) might change due to secular interactions or close encounters. This can be used to narrow down possible orbits of the test planet (Kopparapu et al. 2008; Raymond et al. 2008). The 'fraction of time on detected orbits' (FTD) quantifies the probability that the inner and outer planet are located at their observed best-fitted orbits, inside of the observational error bars. The back-reaction of the detected planets might be strong enough so that they spend a significant time outside the (a, e)-region they are observed in. Hence, the smaller the FTD the more unlikely the presence of a hypothetical planet on the corresponding initial orbit. This method is only applicable to systems were the secular interactions between the detected planets are small. Otherwise, the osculating elements a and e of the detected planet oscillate beyond their accredited orbits periodically without the influence of a hypothetical planet (Veras & Ford 2009). Hence, we do not apply the FTD as an absolute constrain and only use it for planets which do not leave their observed (a, e)-region despite secular interaction with other observed planets in the system.

Secular interaction among planets is a well-studied field. The Lagrange–Laplace secular evolution theory, well described in Murray & Dermott (2000), allows us to predict the long-term evolution of eccentricity and inclination in multiplanet systems. The secular perturbation of the orbital elements are then given by the disturbing function expanded to second order in eccentricity and

inclination. Thus, this classical theory demands that eccentricities and inclinations are small enough to guarantee that such an expansion is adequate. While all our simulations start with a small inclination, the eccentricities are sometimes rather large. However, since we use secular theory only as qualitative guideline to check the simulation results, it is not necessary to use higher order secular solutions (e.g. Veras & Armitage 2007).

Here, we apply the secular theory to calculate the effect of a known two-planet system on the hypothetical (massless) super-Earth, following Adams & Laughlin (2006). This holds for a massless particle, but it might hold also for super-Earths, since the known planets in the systems are often much larger. With secular theory, the forced eccentricity component of a test particle can be calculated as a function of semimajor axis and time. Secular theory shows that the osculating eccentricity e of a particle is composed of the time-dependent forced eccentricity $e_{\rm forced}$ and the free eccentricity $e_{\rm free}$ (Murray & Dermott 2000). While the forced eccentricity is caused by the secular interactions with the known planets, the free eccentricity is basically given by the boundary conditions. The maximum value of e is given by $e_{\rm forced} + e_{\rm free}$, the minimum is given by $|e_{\rm forced} - e_{\rm free}|$. If $e_{\rm forced} > e_{\rm free}$, particle oscillates around $e_{\rm forced}$ with amplitude $e_{\rm free}$. Otherwise, it oscillates around $e_{\rm free}$ with amplitude $e_{\rm forced}$.

Most of the systems we study harbour planets on non-circular orbits. As mentioned above, secular interactions will force the orbits of neighbouring test planets to change in eccentricity. On the other hand, MMRs or close encounters can cause a change in semimajor axis. To record the actual location of the test planets during the simulations, the (a, e)-plane is divided in multiple bins. The number of massive planets located in this bin in all simulations is summed over all time steps. Binning the presence of a stable particle in the (a, e)-plane results in the time-averaged location of all particles. It reveals the most likely eccentricity of a hypothetical planet for a given semimajor axis when it will be observed.

4 RESULTS

The massless test particle simulations reveal that not all systems are worth further detailed study. They show that the HIP 14810 triple giant plant system harbours test particles in a well-defined region in between the two inner planets. Between the two outer planets, where the HZ is located, only very few orbits are stable. Hence, we do not perform additional simulations with massive test planets. HD 37124 hosts three giant planets of almost equal mass. The inner edge of the HZ coincides with the apocentre of the inner most planet. As a result of the relatively high masses of the planets and their non-zero eccentricity, all test particles are lost in a few 100 000 yr and we do not carry out the simulations with massive test planets.

Finally, we focus on the eight systems that are most likely to provide stable orbits in the EHZ. Hence, we carry out massive test planet simulations for the systems HD 11964, HD 47186, HD 147018, HD 163607, HD 187123, HD 190360, HD 217107 and HIP 57274. The main results are given in Table 3 and Figs 2 and 3. They show the location of the stable orbits of 10 $M_\oplus$ mass planets in the systems, given that the orbital solution for the known planets is correct. HD 168443 hosts two known companions: the inner one is a very massive giant planet, the outer a brown dwarf. Test particle simulations reveal that some stable orbit exist around 1 au at low eccentricities. Although this is not part of the EHZ, we carry out the massive test planet simulations to check if massive super-Earths may survive. Only very few planets are stable. Hence, this system is not shown as a figure.

In Table 3 the fractions of orbits that are stable ($f_{\rm stab}$) are listed. The normalized fractions $F_{\rm stab}$ are given by the percentage of the area in $a_{\rm inner} < a < a_{\rm outer}$ and $0.0 < e < 1.0$ that is covered by the stable orbits.

Based on the numerous massive test planet simulations that are carried out, we present the major insights in the following subsection.

4.1 Zones of dynamical influence

The lines of crossing orbits give a fundamental constraint on the stability regions. In addition to the physical cross-section, dynamical interaction plays a major role. In HD 11964, HD 47186, HD 187123, HD 190360 and HD 217107 the shape of the stability region is clearly following the lines of crossing orbits with a partially significant offset. Towards the inner planet, the line traces the outer edge very well, apart from high e. The outer edge of the stability zone is shifted inwards due to the dynamical influence of the outer planet. The relatively large semimajor axis of this planet results in a large Hill radius and dynamical influence. In addition, higher eccentricity of the outer planet leads to a more diffuse transit from stability to instability, in our examples often in combination with MMRs. In contrast, low eccentricity results in a sharp edge (HD 11964).

In the case of HIP 57274, HD 163607 and HD 147018, which are systems with strong interaction among the planets, the stability regions are significantly truncated compared to the line of crossing orbits. Beside the large masses of the inner planets, their relatively large semimajor axes enhance their dynamical influence. In addition, the dominant MMRs of the outer planets amplify this effect. Secular resonances result in the oscillation of the eccentricity of the known planets. Therefore, the lines of crossing orbits change on a secular time-scale. Nevertheless, the shift of the lines of crossing orbits due to the oscillations is too small to truncate the stable region additionally over time.

4.2 Significant MMRs

The MMRs play a major role in shaping the stable regions of many systems. In most of the systems the outer planet has a relatively high eccentricity ($e > 0.1$) and mass ($m > 1.5\,M_{\rm Jupiter}$). The MMRs of this planet with the test planet tend to destabilize the latter. This cuts narrow wedges into the outer part of stability zone (e.g. HD 187123, HD 190360 or HD 217107). Since they can be very narrow, their visibility is sometimes limited by the finite resolution of our sampling. In combination with a highly eccentric ($e > 0.4$) and massive inner planet, the stability region tends to be completely divided by MMRs, because the orbit of the inner planet truncates the stable zone significantly.

In contrast, small planets ($m < 1.0\,M_{\rm Jupiter}$) with low eccentricity ($e < 0.05$) can provide additional stable zones due to MMRs, because their dynamical influence is not that strong. In the region beyond the limits of crossing orbits, test planets can be captured by the strong MMRs (HD 11964:2d:1c and 3d:2c, HD 47186:2d:1c). These MMRs tend to catch particles which would be potentially unstable. In HD 11964, the high-order MMRs (3d:1c) are not strong enough to cut into the stable zone. To exclude a resolution effect, a high resolution zoom in the parameter space around the 2d:1c and 3d:1c MMRs with 8000 simulations was calculated and is shown in

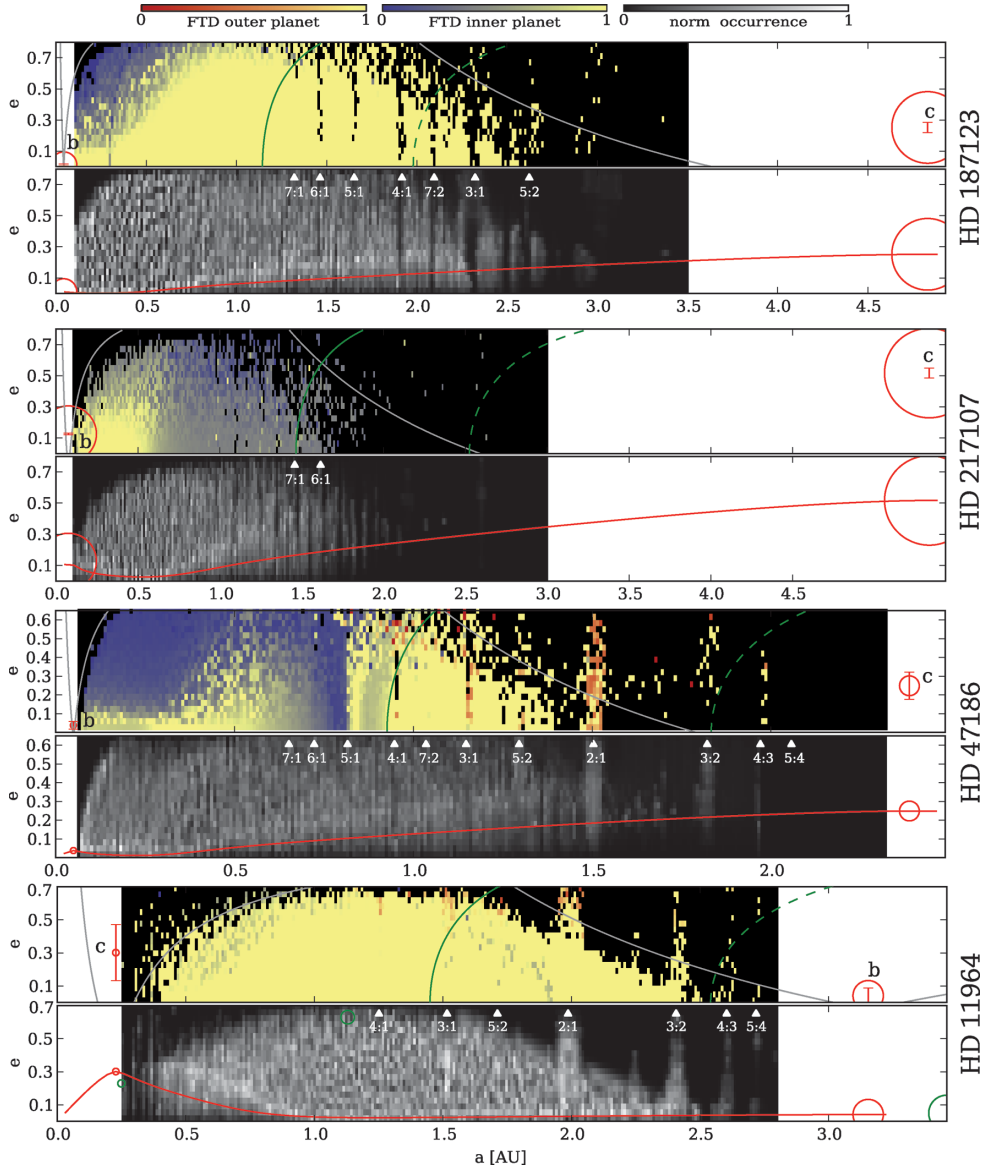

Figure 2. Results of the massive test planet simulations in the systems HD 187123, HD 217107, HD 47186 and HD 11964 with deceasing $\beta/\beta_{\rm crit}$ = (15.09,8.94,6.13,2.04). For each system, the results are presented in two panels. Top panel: the yellow region represents the orbital elements of massive test planets which were stable for 10 Myr. The black regions represent unstable regions. The colour gradient from yellow to red represents orbits with a strong interaction with the inner planet, this means the fraction of time on detected orbit (FTD) decreases. The gradient from yellow to blue represents orbits with a strong interaction with the outer planet (here planet *d*). The grey lines show the location of the crossing orbit of the planets. The full green lines gives the inner edge of the EHZ, the dashed green line gives the outer edge. Bottom panel: the occurrence of a test planet in a given parameter space bin during the whole simulation normalized to 1. The brighter the colour the more likely is it to observe a planet with orbital elements according to this bin. The red line gives the value of the forced eccentricity due to secular perturbation. The location of the most important MMRs is also shown. The green circles in the lower panel of HD 11964 show the three planet solution by Gregory (2007).

Fig. 4. The location of the 3*d*:1*b* and 5*d*:2*b* MMRs in the massless particle simulations are cleaned, while in the massive planet simulations, the orbits captured in the MMRs are stable. A decrease in the FTD of planet *b* at the MMRs indicates a strong interaction among the planets in resonance.

4.3 Fraction of time on detected orbits

The FTD is diminished in large regions of the stability zone in many systems. The inner planet is often perturbed significantly by the test planets. Typically, test planets close to the inner planet play a major role. The more the initial eccentricity coincides with the initial eccentricity of the inner planet, the smaller is their effect (e.g. HD 47186, HD 187123, HD 190360). If the eccentricities do not coincide, the eccentricity of the inner planet is forced to change. MMRs are the only occasions where the FTD of the outer planet is diminished (HD 11964, HD 47186). In this case, the resonance with the test planet is strong enough to perturb the outer planet with $m < 1\,M_{\rm Jupiter}$ significantly. In the zoom simulation of two details of HD 11964 (Figs 4 and 5), this effect is clearly visible.

An interesting feature can be observed in some of the systems: The FTD of the inner planet has a minimum in parts of the stability region while between this minimum and the inner planet, at the same eccentricity, the FTD reaches 1. This is observed in HD 47186, HD

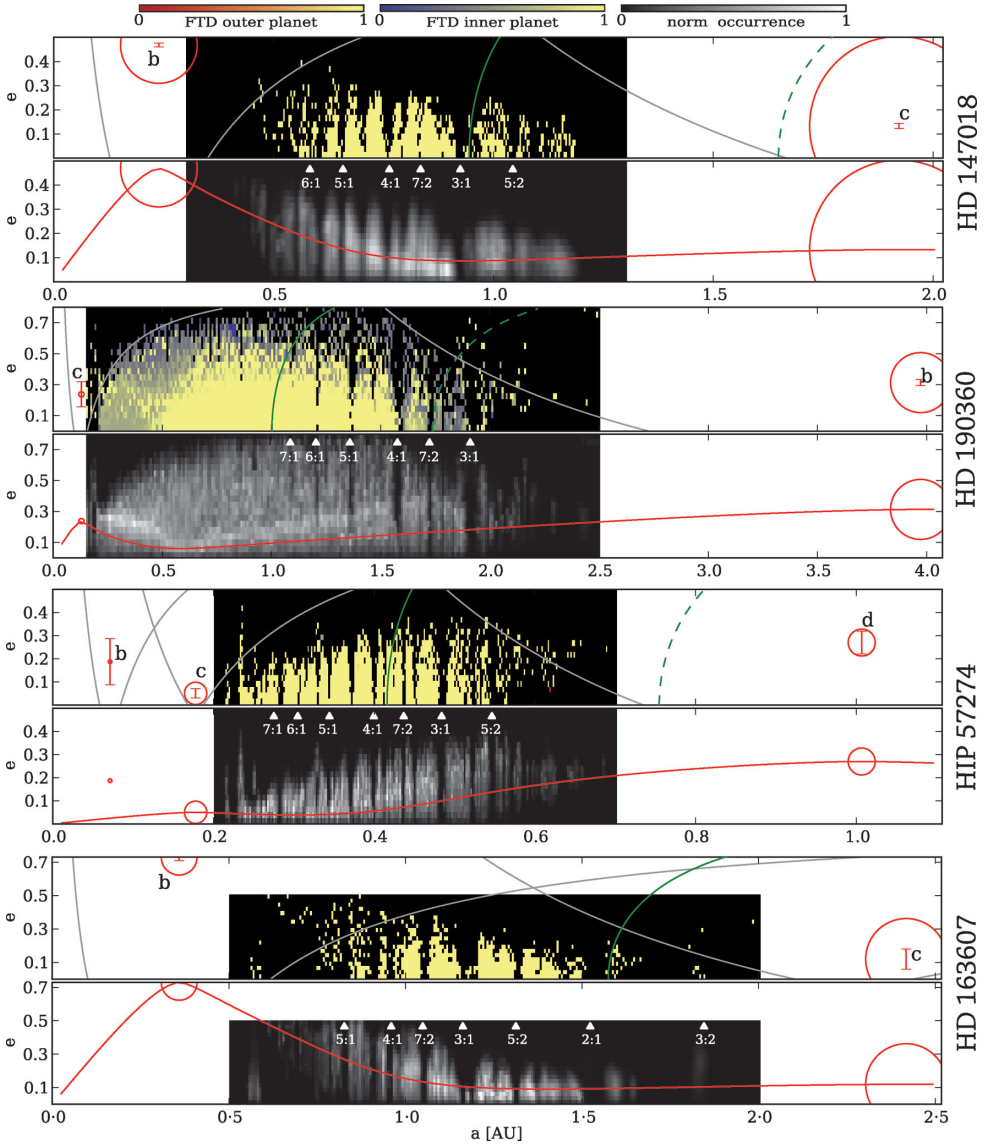

Figure 3. Results of the massive test planet simulations in the systems HD 147108, HD 190360, HIP 57274 and HD 163607 with deceasing $\beta/\beta_{\rm crit}$ = (2.01,1.80,1.78,1.58). For each system, the results are presented in two panels. Top panel: the yellow region represents the orbital elements of massive test planets which were stable for 10 Myr. The black regions represent unstable regions. The colour gradient from yellow to red represents orbits with a strong interaction with the inner planet, this means the fraction of time on detected orbit (FTD) decreases. The gradient from yellow to blue represents orbits with a strong interaction with the outer planet (here planet *d*). The grey lines show the location of the crossing orbit of the planets. The full green lines gives the inner edge of the EHZ, the dashed green line gives the outer edge. Bottom panel: the occurrence of a test planet in a given parameter space bin during the whole simulation normalized to 1. The brighter the colour the more likely is it to observe a planet with orbital elements according to this bin. The red line gives the value of the forced eccentricity due to secular perturbation. The location of the most important MMRs is also shown.

187123, HD 190360, HD 217107 and marginally in HD 11964. This effect is caused by secular resonances and depends very much on the architecture of the system and on the given error bars. An illustrative example is given by HD 47186. The simulations show that there is no continuous region of high FTD at $a < 0.9$ au. In fact, a minimum in the FTD around 0.7 au of FTD ≈ 0.3 with respect to the inner planet is found for all eccentricities. Secular perturbations of the outer planet let the test planet oscillate according to the corresponding $e_{\rm free}$ and $e_{\rm forced}$. This results in the eccentricity oscillation of the inner planet which reacts significantly due to its relatively small mass of around 22 $M_\oplus$. One can say that the test planet acts to transfer a secular perturbation from the outer planet on to the inner one. If the test planet is located closer to the inner planet, $e_{\rm forced}$ is smaller. Therefore, its secular oscillation is too small to affect the inner planets FTD. If the test planet is further away from the inner planet, its forced oscillation can hardly be transferred to the inner planet.

HD 47186 was already studied in detail with lower resolution by Kopparapu et al. (2008). Our stability region agrees with their results, but the FTD results differ. Kopparapu et al. (2008) found a sharp border in the FTD at $a \approx 0.25$ dividing a region of very low ($\approx$0.2) FTD and the broad region of FTD = 1 between 0.3 and 1.3 au. We found out that this disagreement with Kopparapu et al. (2008) is caused by different time steps used in the integration. Secular oscillations of the planets' eccentricities are sometimes missed in (Kopparapu et al., private communication).

Regarding the existence of possible orbits in the EHZ, the FTD provides significant constrains only in the case of HD 217107. This

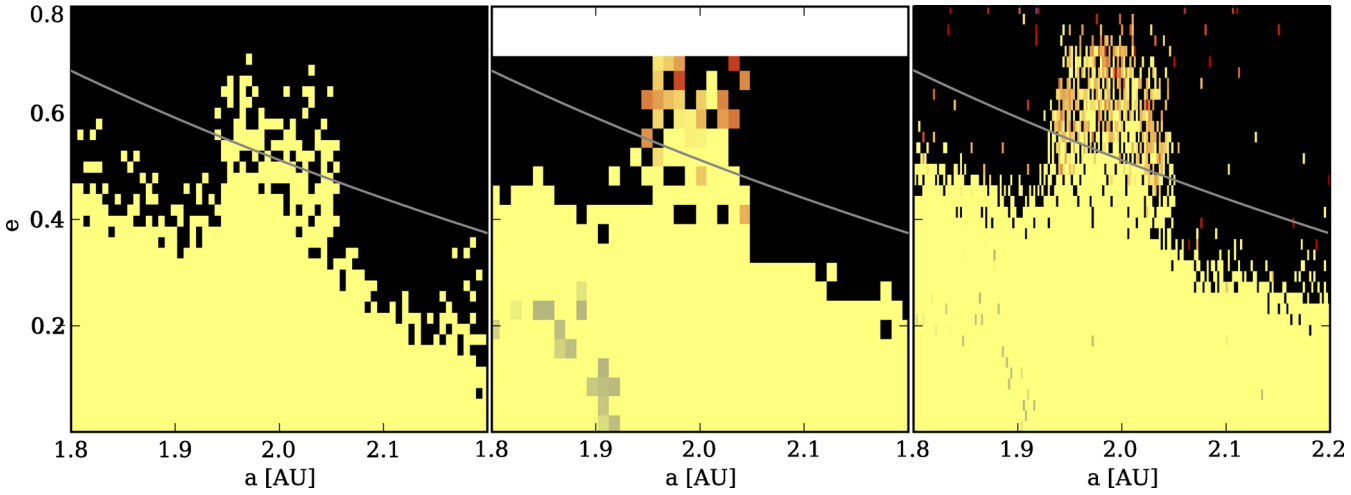

Figure 4. 2*d*:1*c* MMR of the HD 11964 system. The left-hand panel shows a detail of the test planet simulation. The central panel shows a detail of the massive test planet simulations presented in Fig. 2 ($e \leq 0.7$). The right-hand panel shows massive test planets simulations carried out in higher resolution (200×40 simulations). The colour gradient is given in Fig. 2. Particles and test planets initially located close to the resonance (±0.05 au) become stable. If they are located above the line of crossing orbits, they significantly diminish the FTD of planet *b*.

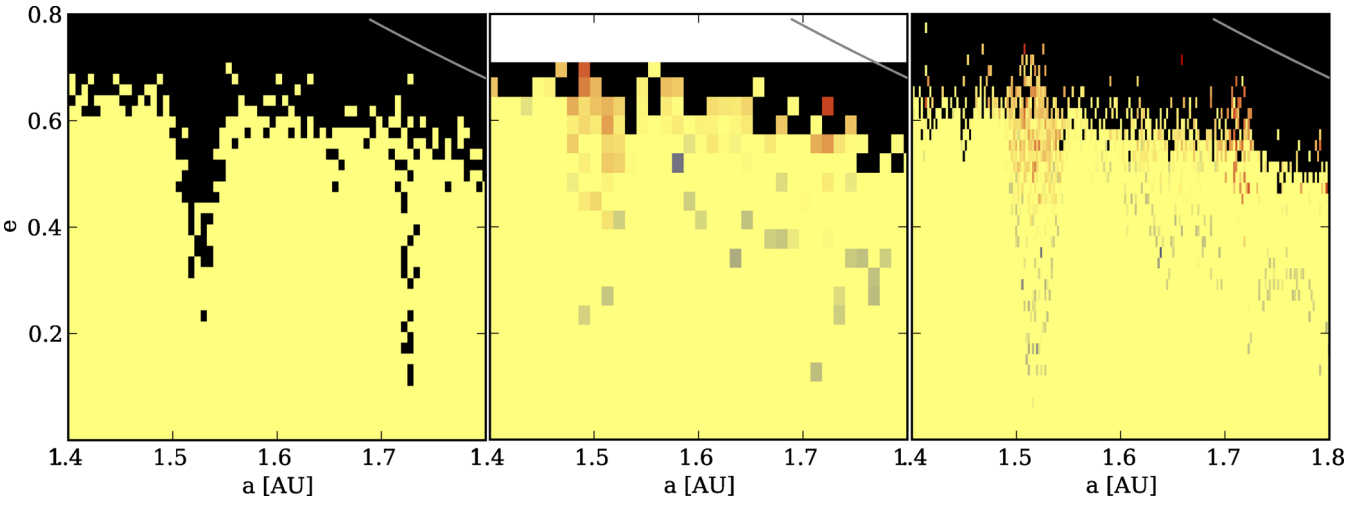

Figure 5. 3*d*:1*c* and 5*d*:2*c* MMRs of the HD 11964. The left-hand panel shows a detail of the test planet simulation. The central panel shows a detail of the massive test planet simulations presented in Fig. 2 ($e \leq 0.7$). The right-hand panel shows massive test planets simulations carried out in higher resolution (200×40 simulations). While massless particle in MMR with the outer planet become unstable, the massive test planets are stable. They diminish mostly the FTD of the outer planet.

is a result of the average location of the HZ, whose distance to the inner planet is often large and resulting secular perturbations are small.

The FTDs of the known planets were not studied in the case of the planets in HD 163607 and HD 147186 and planet *c* in HIP 57274. They were excluded because of strong secular perturbations among the known planets.

4.4 Massless test particles

The massless test particle simulations provide very detailed pictures of the stability regions. Comparisons of the area of the stable zone found in the massless test particle simulations and results of the massive test planet simulations show that both are very similar. The normalized percentages of stable orbits are listed in Table 3. The most significant difference is prominently seen in HD 11964. In Fig. 4, a detailed comparison with the test particle simulation, the low resolution and the high resolution simulation set of massive test planets is shown. The location of the 4*d*:1*b* and 3*d*: 1*b* MMRs in the massless particle simulations are cleaned, whereas in the massive planet simulations, the planets in the MMRS are stable. Obviously, the mass of the test planet adds additional stability to the MMRs. Beside the MMRs, the low and high resolution simulations with massive test planets agree very well with the massless particle simulations. In some parts, it seems that the test particle simulation cannot reproduce the complete area of stable orbits at the very edge of the stability region. Beside the effect of lower resolution, a possible explanation is that test planets involved in close encounters are not as much perturbed as massless particles.

4.5 Forced eccentricity

The lower panel of each system's plot (Figs 2 & 3) shows the normalized occurrence rate. It gives the time-averaged location of all

stable orbits and represents the likelihood that a hypothetical planet is found in a certain bin of the (a,e)-plane. Many systems show a prominent curve of maximum occurrence rate (e.g. HD 190360, HD 47186). The curves approach asymptotically the eccentricity of the inner and outer planets, often with a minimum in eccentricity. This shows that the test planets are forced to change their eccentricity.

When comparing the analytically estimated amplitude of $e_{\rm forced}$ given by secular theory and the most likely location of the test planets in the (a,e)-plane, the minimum of the predicted forced eccentricity clearly coincides with the minimum of the curve in the occurrence rate. The maximum of the occurrence rate along the eccentricity does not agree with $e_{\rm forced}$. Beside the limitations of the secular theory at high eccentricities, this is caused by the fact that the particle oscillates around $e_{\rm forced}$ only when $e_{\rm forced} > e_{\rm free}$. Hence, the planets with initially high eccentricity tend to spend most of their time at high e. Therefore, we have to point out that the measured occurrence rate depends on the expansion of the sampled region along the e-axis.

Although the averaged flux that the planet receives is more important for habitability than the planet's eccentricity (appendix A), a small $e_{\rm forced}$ can be interpreted as an optimal location for a habitable planet, following Adams & Laughlin (2006). When we assume that the particle is initially on a low eccentric orbit, the $e_{\rm forced}$ gives the more realistic eccentricity than the occurrence rate.

4.6 Close encounters

The numbers of stable test planets that were part of a close encounter are given in Table 3. The fraction of such stable orbits is ≈15 per cent in HD 147018 or ≈4 per cent in HD 11964. (In HD 168443, almost all stable test planets had a close encounter but since only ≈0.1 per cent of all configurations are stable, this is not surprising.) Most close encounters take place at the outer edge of the stability region. This confirms our decision not to classify an orbit as unstable as soon as its planet has a close encounter. Thus, the criterion to identify unstable orbits should not be given by the occurrence of a close encounters. Nevertheless, there are systems where no close encounters of the stable test planets take place.

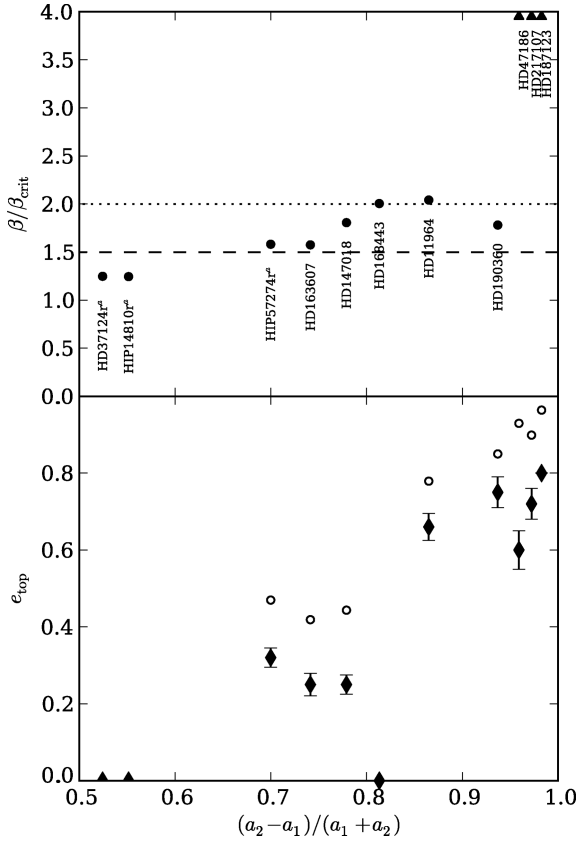

Figure 6. Constraining stability zones analytically. Various measures of stability shown as a function of the normalized spacing of the enclosing planets. Systems with more than two known planets are marked with superscript a. Top panel: analytic stability criterion $\beta/\beta_{\rm crit}$. The dashed lines indicate the minimum value for a system to enclose additional planets. Depending on the planet configuration, this line can shift up to 2.0 (dotted line). System above the dotted line always allow stable orbits. Bottom panel: the maximum eccentricity $e_{\rm top}$. It is shown as circles when c_i are obtained by fitting a piecewise curve to the data. The directly measured $e_{\rm top}$ from the simulations and their uncertainties are given as diamonds.

4.7 Analytic predictions

The top panel of Fig. 6 shows $\beta/\beta_{\rm crit}$ of the planetary systems. In the case of the two systems that have the smallest separation in semimajor axis, they are well below $\beta/\beta_{\rm crit} < 1.5$. For the most separated systems, $\beta/\beta_{\rm crit} > 2.0$. In between, there are systems with $1.5 < \beta/\beta_{\rm crit} < 2.0$ where the existence of additional enclosed stable orbits is not sure. The simulations show that in our sample, all system with $\beta/\beta_{\rm crit} > 1.5$ can harbour additional super-Earths. Nevertheless, HD 168443 is right at the edge of $\beta/\beta_{\rm crit} = 2.0$ and only very few planets are stable.

The bottom panel of Fig. 6 shows the maximum eccentricity $e_{\rm top}$ as a function of the separation. Systems containing planets with zero eccentricity would follow a straight line (e.g. Fang & Margot 2012) whereas high eccentricity planets with high masses are truncating the stable region, respectively $e_{\rm top}$, or even allow no stable region ($e_{\rm top} \leq 0$). Large orbital spacing of the planets suppresses this effect. We estimate c_1 and c_2 for every system separately. Then, calculating $e_{\rm top}$ results in a slight overestimation with respect to the maximum eccentricity observed directly in the simulations, because the piecewise function does not account correctly for the flatted top of the stable region. Hence, even if we would guess c_1 and c_2 correctly, we would overestimate slightly the height of the stability zone with this analytic approach.

4.8 Predicting habitable super-Earths

HD 168443 provides only very few stable simulations. Hence, we treat it as a fully packed system. All the systems we study in detail with massive test planets provide well-defined regions with stable orbits for a $10\,M_\oplus$ super-Earth, partially located in the EHZ. We combine the stability of the orbits with the time-averaged location given by the occurrence rate, the analytic estimation of $e_{\rm forced}$ and the weak constraints from the FTD values. The location in the (a, e)-plane where a hypothetical super-Earth is most likely to be observed is given in Table 4 for each system.

There exist predictions from previous studies. HD 47186 was studied in detail concerning the possible existence of a planet in the EHZ by Kopparapu et al. (2008). They found that a $10\,M_\oplus$ planet is stable in the EHZ or even two $10\,M_\oplus$ with low eccentricities can exist between planets b and c. As mentioned above, we give a different estimate of the FTD map. The differences result from larger time steps used in the Kopparapu et al. (2008) simulations.

Table 4. The most likely location in the (a,e)-plane for the observation of a hypothetical habitable super-Earth. We comment on the system if there are features that could limit the habitability (high e) or the stability (small FTD, MMRs).

System	Stable region in HZ (a, e)	Comment
HD 11964	(1.3–2.4 au, 0.05)	–
HD 47186	(0.9–1.3 au, 0.1–0.3)	High e
HD 147018	(0.8–0.9 au, 0.0–0.1)	–
HD 163607	(1.3–1.4 au, 0.05–0.1)	–
HD 187123	(1.0–2.2 au, 0.1–0.3)	High e
HD 190360	(0.8–1.5 au, 0.1–0.3)	High e
HD 217107	(1.3–1.6 au, 0.3)	Small FTD
HIP 57274	(0.37–0.56 au, 0.1–0.3)	Strong MMRs

Gregory (2007) proposed that the planetary system HD 11964 consists of three instead of two planets based on fitting the Doppler spectroscopy data. Their three-planet solution is shown as green circles in Fig. 2 and is consistent with our stability region. The small difference in the orbital elements of the known planets would not significantly change the region. Nevertheless, the high eccentricity of the additional planet seems very unlikely and is outside of the EHZ. This three-planet solution was not confirmed by Wright et al. (2009).

In HD 190360, Veras & Ford (2010) reported a stable terrestrial planet in the HZ might be possible according to test particle stability simulations, in agreement with our results.

In Jones et al. (2006), numerous systems are studied and the stability of a habitable Earth is estimated based on critical distances (basically parametrized by c_1 and c_2, see Section 2.2) to the giant planets. Hence, their estimation of the stability zone differs fundamentally from our fully numerical approach. Since in some system new planets were found in the meantime, we can only compare our results concerning HD 190360, HD 168443, HD 217107 and HD 37124. We agree on the survival of hypothetical planets in the HZ of HD 190360. We also found that stable orbits are unlikely in the HZ of HD 168443. In HD 217107, Jones et al. (2006) localized the HZ at $2.0 \lesssim a \lesssim 4.0$, which differs from our estimate. This results from the fact that we use slightly different stellar parameters and a newer estimation of the HZ (Kopparapu et al. 2013). According to our results, the HZ is closer to the star and thus, the stability zone is partially located inside the HZ. In addition, Jones et al. (2006) predicted that stable orbits can exist partially in the HZ of HD 37124. Out test particle simulations show that no additional planets are stable in the HZ and the analytic approach fails in this system.

4.9 Limited parameter space

There are many parameters that control the architecture of a 2+1 planet system. Our simulations focus only on two dimensions (semi-major axis and eccentricity) of a multidimensional parameter space. Orbital inclinations, orbital phases and the mass of the test planet offer a wide range of additional scenarios to study. We think that only extreme values in inclination and mass will have a significant effect on the results: the orbital angles of the planets were chosen randomly in the simulations and only at the edges of the stable zone do these angles play any role regarding stability or FTD values. This could explain why the FTD does not always have a continuous gradient; meaning that sometimes small FTD values alternate with FTD $\approx$1 at the transition from high FTD to low FTD regions (for example HD 47186, $a \approx 0.5$, $e > 0.2$). Since massless and massive test planet simulations give very similar results, only test planets with masses $m \gg 10\,\mathrm{M}_\oplus$, small Neptune's, might put the stability of the system at risk. Beside the parameters that control the orbit of the hypothetical planet, the orbital solution of the known planets is not unique. High inclination and masses can have a dramatic effect on the stability zone or on the stability of the known planets (Veras & Ford 2010).

Our simulations show that there are broad stable regions in many of the systems. These regions can harbour more than one super-Earth size planet. But the parameter space increases rapidly by adding new planets, and we did not take this into account in additional simulations.

To test if the significance of our results depends on the simulation period of 10 Myr, we carried out the simulations of HD 190360 for 50 Myr. Indeed, we observed that the fraction of stable orbits reduces from 48.2 to 46.3 per cent. The overall shape and extension of the stability zone is not affected. Mostly, the additional unstable orbits are located at the MMRs which tend to stabilize orbits and the MMR are a bit more pronounced. All told, the limitation to 10 Myr seems reasonable and does not influence our final results.

5 SUMMARY AND CONCLUSIONS

We carry out numerous N-body simulations with the new GPU code GENGA to study the existence of hypothetical planets in extra-solar planetary systems. In nine systems, we study the stability of a 10 $\mathrm{M}_\oplus$ super-Earth in high resolution in the (a, e)-plane. The reaction of the known planet on this hypothetical body and its movement in the (a, e)-parameter space allow us to predict the most likely orbital parameters of a super-Earth in the HZ. Following the PPS-hypothesis, we find that for eight systems additional low-mass planets can exist (apart from HD 168443), most of them with possible orbits in the EHZ (apart from HD 217107). The most promising candidate hosting a stable super-Earth in its HZ with low e is HD 11964 and, with a modest eccentricity of $e \approx 0.2$: HD 47186, HD 187123 and HD 190360.

Beside the lines of crossing orbits, MMRs with the outer planet are a main feature that shaped the stable region. Comparing the simulations with massless test particles and the simulations with massive planets, the main differences are found in the effect of the MMRs. While the 3 : 1 MMR in HD 11964 results in an unstable wedge in the stable region, the same MMR is stable if the hypothetical planet is massive.

Simulations in several systems show that close encounters are not a good criteria to identify unstable orbits. In some systems, a significant fraction of the planets on stable orbits are involved in such an energy exchange.

Beside the drawbacks of the FTD values, it does not constrain any of our stable zones in the HZ significantly (apart from HD 217107). We point out that our FTD results concerning HD 47186 are fundamentally different to a previous study and shows some interesting secular resonance effects.

ACKNOWLEDGEMENTS

We acknowledge the support of the Swiss National Science Foundation Grant No. 20020-127896 and thank the University of Zurich and HP^{2}C project for the financial support. The code was developed at CSCS and with HP^{2}C resources. The simulations were performed on zBox 4 supercomputer on NVIDIA GTX 590 graphic processor units at the University of Zurich. Thanks for the technical support to Doug Potter. We thank an anonymous reviewer for his helpful suggestions as well as Ravi Kopparapu and Ben Moore for

useful discussions. This research has made use of the Exoplanet Orbit Database and the Exoplanet Data Explorer at exoplanets.org.

REFERENCES

Adams F. C., Laughlin G., 2006, ApJ, 649, 992
Asghari N. et al., 2004, A&A, 426, 353
Baluev R. V., 2009, MNRAS, 393, 969
Barnes R., Greenberg R., 2006, ApJ, 647, L163
Barnes R., Greenberg R., 2007, ApJ, 665, L67
Barnes R., Quinn T., 2004, ApJ, 611, 494
Barnes R., Raymond S. N., 2004, ApJ, 617, 569
Barnes R., Raymond S. N., Jackson B., Greenberg R., 2008, Astrobiology, 8, 557
Bean J. L., McArthur B. E., Benedict G. F., Armstrong A., 2008, ApJ, 672, 1202
Borucki W. J. et al., 2011, ApJ, 728, 117
Borucki W. J. et al., 2012, ApJ, 745, 120
Butler R. P., Marcy G. W., Fischer D. A., Brown T. M., Contos A. R., Korzennik S. G., Nisenson P., Noyes R. W., 1999, ApJ, 526, 916
Chambers J. E., 1999, MNRAS, 304, 793
Dumusque X. et al., 2012, Nat, 491, 207
Fang J., Margot J.-L., 2012, ApJ, 751, 23
Fischer D. A. et al., 2008, ApJ, 675, 790
Fressin F. et al., 2012, Nat, 482, 195
Gladman B., 1993, Icarus, 106, 247
Gregory P. C., 2007, MNRAS, 381, 1607
Hamilton D. P., Burns J. A., 1992, Icarus, 96, 43
Hinse T. C., Michelsen R., Jørgensen U. G., Goździewski K., Mikkola S., 2008, A&A, 488, 1133
Jones B. W., Sleep P. N., Underwood D. R., 2006, ApJ, 649, 1010
Kasting J. F., Whitmire D. P., Reynolds R. T., 1993, Icarus, 101, 108
Kopparapu R. K., Hanna C., Kalogera V., O'Shaughnessy R., González G., Brady P. R., Fairhurst S., 2008, ApJ, 675, 1459
Kopparapu R. K. et al., 2013, ApJ, 765, 131
Lissauer J. J. et al., 2011, Nat, 470, 53
Lo Curto G. et al., 2013, A&A, 551, A59
Mandell A. M., Sigurdsson S., 2003, ApJ, 599, L111
Marchal C., Bozis G., 1982, Celest. Mech., 26, 311
Matsumura S., Ida S., Nagasawa M., 2013, ApJ, 767, 129
Menou K., Tabachnik S., 2003, ApJ, 583, 473
Murray C. D., Dermott S. F., 2000, Sol. Sys. Dyn.
Pepe F. et al., 2011, A&A, 534, A58
Raymond S. N., Barnes R., 2005, ApJ, 619, 549
Raymond S. N., Mandell A. M., Sigurdsson S., 2006, Sci, 313, 1413
Raymond S. N., Barnes R., Gorelick N., 2008, ApJ, 689, 478
Saha P., Tremaine S., 1992, AJ, 104, 1633
Schneider J., Dedieu C., Le Sidaner P., Savalle R., Zolotukhin I., 2011, A&A, 532, A79
Tuomi M., Anglada-Escudé G., Gerlach E., Jones H. R. A., Reiners A., Rivera E. J., Vogt S. S., Butler R. P., 2013, A&A, 549, A48
Veras D., Armitage P. J., 2007, ApJ, 661, 1311
Veras D., Ford E. B., 2009, ApJ, 690, L1
Veras D., Ford E. B., 2010, ApJ, 715, 803
Vogt S. S., Butler R. P., Haghighipour N., 2012, Astron. Nachr., 333, 561
Williams D. M., Pollard D., 2002, Int. J. Astrobiology, 1, 61
Wisdom J., Holman M., 1991, AJ, 102, 1528
Wittenmyer R. A., Endl M., Cochran W. D., Levison H. F., Henry G. W., 2009, ApJS, 182, 97
Wright J. T., 2010, in Gożdziewski K., Niedzielski A., Schneider J., eds, EAS Publications Series, Vol. 42 Extrasolar planets in multi-body systems: theory and observations. EDP Sciences, Les Ulis Cedex, France, p. 3
Wright J. T., Upadhyay S., Marcy G. W., Fischer D. A., Ford E. B., Johnson J. A., 2009, ApJ, 693, 1084
Wright J. T. et al., 2011, PASP, 123, 412
Yoshida H., 1991, in Kinoshita H., Yoshida H., eds, 24th Symposium on Celestial Mechanics. Tokyo, Japan, p. 132

APPENDIX A: HABITABLE ZONE

Kasting, Whitmire & Reynolds (1993) (recently updated by Kopparapu et al. 2013) provide the inner and outer boundaries of the HZ. A planet on an eccentric orbit may partially escape from the HZ, even if its semimajor axis lies inside the HZ. Williams & Pollard (2002) showed that orbit-average flux is the most important parameter for a long-term climate stability. The boundaries of the HZ around a star depends on its luminosity L and its effective temperature $T_{\rm eff}$ as well as on planetary characteristics that control the greenhouse effect. The flux depends mainly on the luminosity. The effective temperature is a measure of the infrared fraction in L. A greater infrared fraction results in a greater greenhouse effect for a given stellar flux. Following the new estimates Kopparapu et al. (2013), the critical flux at the inner boundary of the HZ, where runaway greenhouse effect would take place and all surface water will evaporate and hydrogen will rapidly escape to space, is given by

$$S_i = 1.0140 + 8.1774 \times 10^{-5} T_* + 1.7063 \times 10^{-9} T_*^2 - 4.3241 \times 10^{-12} T_*^3 - 6.6462 \times 10^{-16} T_*^4, \quad \text{(A1)}$$

where $T_* = T_{\rm eff} - 5740$ K. The outer boundary flux corresponds to a minimum flux at which a maximum greenhouse effect can maintain liquid water on the surface of the planet with a cloud-free carbon dioxide atmosphere,

$$S_o = 0.3483 + 5.8942 \times 10^{-5} T_* + 1.6558 \times 10^{-9} T_*^2 - 3.0045 \times 10^{-12} T_*^3 - 5.2983 \times 10^{-16} T_*^4. \quad \text{(A2)}$$

The critical distances denoting the boundaries of the HZ are then given by the inverse square law:

$$\frac{r_i}{r_{\rm au}} = \left(\frac{1}{S_i} \frac{L_*}{\mathrm{L}_\odot} \right)^{1/2}, \quad \text{(A3)}$$

$$\frac{r_o}{r_{\rm au}} = \left(\frac{1}{S_o} \frac{L_*}{\mathrm{L}_\odot} \right)^{1/2}. \quad \text{(A4)}$$

$\mathrm{L}_\odot$ is the solar luminosity and $L_* = 4\pi R_* \sigma T_{\rm eff}$ is the luminosity of the star, a function of the radius of the star, R_*. $r_{\rm au}$ denotes the distance of Sun and Earth.

We focus on planets which receive as much flux over one orbit as a planet on circular orbit with the same semimajor axis confined in the HZ, we have to take into account the eccentricity dependent orbit-averaged mean flux (Williams & Pollard 2002; Adams & Laughlin 2006):

$$\langle F \rangle = \frac{F}{4\pi a^2 \sqrt{1 - e^2}}. \quad \text{(A5)}$$

Hence, we assume that his flux corresponds to the critical fluxes at the HZ boundaries for $e = 0$ and we can deduce constraints for an orbit with elements (a, e) inside these boundaries:

$$r_i < a(1 - e^2)^{1/4} < r_o. \quad \text{(A6)}$$

We will refer to this concept of the HZ as the EHZ (Barnes et al. 2008; Kopparapu et al. 2008).

This paper has been typeset from a TEX/LATEX file prepared by the author.

6

Prospects

The work presented in this dissertation is a single step on a long way towards understanding planet formation. As we have seen, the papers request further work and improvements: there are uncertainties which affect the estimated occurrence rate of Earth-Moon planetary systems, which could be reduced or eliminated. The elemental composition of terrestrial planets were obtained based on a model that has many caveats. A better understanding of the sources of these caveats would open new possibilities to study the bulk abundances of planets. The presented stability studies could be expanded by more simulations to cover more of the multi-dimensional parameter space as well as to examine other systems. All these points were discussed separately in the particular chapters. Here, we want to highlight which further steps would result in new insights regarding the discussed topics and terrestrial planet formation in general.

Terrestrial planet formation simulations benefit from including smoothed particle hydrodynamic (SPH) simulations or similar methods to determine the collision outcome especially during the giant impact phase. The presented estimation of the Earth-Moon occurrence rate is affected by many sources of uncertainties. Instead of using an semi-analytic interpolation to estimate the circumplanetary disk mass, carrying out SPH simulations directly would clearly improve the results. Furthermore, collisional remnants of giant impacts can cause dynamical friction and thus alter the orbital evolution of planetary embryos. In addition, this debris can form new planetesimals and protoplanets which provide a continuous source for accretion.

The elemental composition of planets is also affected by the outcome of giant impacts. On the one hand, high temperatures generated in collisions cause a loss of volatile material. Subsequent collisions result in a depletion of volatiles and in a enrichment of refractory elements. This is not taken into account in our paper, which becomes manifest in the overestimation of volatiles in the planets. On the other hand, the material ejected by the collisions can be captured by other planets in the system and alter their bulk composition [119]. This is an additional mechanism

for radial mixing and equilibration of the elemental compositions of the planets in the system. However, the amount of exchanged material might be very small. All in all, replacing the perfect merging of particles in N-body simulations by more sophisticated collision outcomes, simulated individually or tabulated, and a suitable treatment of the collisional remnants would be desirable.

The N-body simulations used in the first two papers are limited to 2000 fully interacting particles. Although ahead of previous simulations, the initial bodies have a mass of around half of the Moon and are far away from being the initial planetesimals as described by the theoretical models. To better understand the dependence on the initial conditions and the runaway growth phase, sophisticated simulations with more particles, maybe millions of them [120], would be useful.

The focus of many N-body simulations on the Solar System is both comprehensible and limiting. Other system architectures exist, where Super-Earths or hot Jupiters are present or where no giant planets affect terrestrial planet formation. In such systems, the boundary conditions are different compared to the Solar System. This could apply to the star mass, the profile of the protoplanetary disk, the location of the giants, the disk life time, binary companions to the star and so on. Consequently, many parameters we studied would be altered: the diversity in the composition of the protoplanets, the mechanisms for water delivery, the moon formation rate or even the existence of terrestrial planets.

The surprisingly similar composition of the Earth and the Moon is constraining significantly Moon formation simulations, e.g. [121, 122], as well as the composition of the colliding protoplanets. Hence, the estimation of the bulk composition of protoplanets during the formation process traced via N-body simulations can provide both the initial geometry of the collision and the composition of the colliding bodies. Including debris from previous impacts on the planet or ejected by impacts on other protoplanets would help to get a more detailed picture of the origin of the differences in the elemental composition of the Earth and the Moon.

Studying the emergent spectra of hot molten protoplanets could help to verify the late stages of planet formation. The detectability of terrestrial planets during their formation depends amongst others on the chemical composition of the protoplanet, especially on properties of its initial atmosphere [123]. The next generation of telescopes will provide the possibility to detect such protoplanets. Then, models which provide the abundances of chemical elements in protoplanets would be helpful to connect planet formation theory and observations. Especially the occurrence rate of such hot and observable protoplanets could be compared to predictions. However, a sophisticated treatment of the volatile elements during the collisions of protoplanets is required to draw conclusions about the young planet's atmosphere regarding thickness and composition.

To study the formation of planetary systems as a whole, i.e. the emergence of planetesimals, the formation of giant planets via core accretion and the subsequent formation of terrestrial planets as well as a simultaneous production of dwarf planets, asteroid belts, satellites and so on is a challenging task. It involves processes that take place on a short time scale (e.g. collisions), medium time scale (gas giant core accretion, accretion of the envelope) and long time scale (giant impact phase inside the ice line) as well as processes that span different orders of magnitude in size. Thus,

the range in mass and time scales is huge. This requires to study planet formation stepwise by means of different simulations and models which provide insight in different aspects of planet formation. At the same time, the transitions between the models can cause uncertainties as we have seen in the first as well as in the second paper.

Nevertheless, all-embracing simulations that show the complete planet formation from dust to protoplanets will not come true in the next years. However, ongoing observations, new instruments and better codes and tools are going to answer more and more of the open questions concerning terrestrial planet formation and the properties of habitable planets in extrasolar systems. Without doubt, the worlds beyond the Blue Planet are going to stay an exciting topic in the future.

Bibliography

[1] A. Bouvier and M. Wadhwa. The age of the Solar System redefined by the oldest Pb-Pb age of a meteoritic inclusion. *Nature Geoscience*, 3:637–641, September 2010. doi: 10.1038/ngeo941.

[2] M. Duncan, T. Quinn, and S. Tremaine. The origin of short-period comets. ApJ, 328:L69–L73, May 1988. doi: 10.1086/185162.

[3] J. H. Oort. The structure of the cloud of comets surrounding the Solar System and a hypothesis concerning its origin. Bull. Astron. Inst. Netherlands, 11: 91–110, January 1950.

[4] R. M. Canup and W. R. Ward. Formation of the Galilean Satellites: Conditions of Accretion. AJ, 124:3404–3423, December 2002. doi: 10.1086/344684.

[5] D. Jewitt and N. Haghighipour. Irregular Satellites of the Planets: Products of Capture in the Early Solar System. ARA&A, 45:261–295, September 2007. doi: 10.1146/annurev.astro.44.051905.092459.

[6] M. Touboul, T. Kleine, B. Bourdon, H. Palme, and R. Wieler. Late formation and prolonged differentiation of the Moon inferred from W isotopes in lunar metals. Nature, 450:1206–1209, December 2007. doi: 10.1038/nature06428.

[7] M. Mayor and D. Queloz. A Jupiter-mass companion to a solar-type star. Nature, 378:355–359, November 1995. doi: 10.1038/378355a0.

[8] W. J. Borucki, D. G. Koch, G. Basri, N. Batalha, T. M. Brown, S. T. Bryson, D. Caldwell, J. Christensen-Dalsgaard, W. D. Cochran, E. DeVore, E. W. Dunham, T. N. Gautier, III, J. C. Geary, R. Gilliland, A. Gould, S. B. Howell, J. M. Jenkins, D. W. Latham, J. J. Lissauer, G. W. Marcy, J. Rowe, D. Sasselov, A. Boss, D. Charbonneau, D. Ciardi, L. Doyle, A. K. Dupree, E. B. Ford, J. Fortney, M. J. Holman, S. Seager, J. H. Steffen, J. Tarter, W. F. Welsh, C. Allen, L. A. Buchhave, J. L. Christiansen, B. D. Clarke, S. Das, J.-M. Désert, M. Endl, D. Fabrycky, F. Fressin, M. Haas, E. Horch, A. Howard, H. Isaacson, H. Kjeldsen, J. Kolodziejczak, C. Kulesa, J. Li, P. W. Lucas, P. Machalek, D. McCarthy, P. MacQueen, S. Meibom, T. Miquel, A. Prsa, S. N. Quinn, E. V. Quintana, D. Ragozzine, W. Sherry, A. Shporer, P. Tenenbaum, G. Torres, J. D. Twicken, J. Van Cleve, L. Walkowicz, F. C. Witteborn, and M. Still. Characteristics of Planetary Candidates Observed by Kepler. II. Analysis of the First Four Months of Data. ApJ, 736:19, July 2011. doi: 10.1088/0004-637X/736/1/19.

[9] R. P. Butler, G. W. Marcy, E. Williams, C. McCarthy, P. Dosanjh, and S. S. Vogt. Attaining Doppler Precision of 3 M s-1. PASP, 108:500, June 1996. doi: 10.1086/133755.

[10] C. Lovis, F. Pepe, F. Bouchy, G. Lo Curto, M. Mayor, L. Pasquini, D. Queloz, G. Rupprecht, S. Udry, and S. Zucker. The exoplanet hunter HARPS: unequalled accuracy and perspectives toward 1 cm s^{-1} precision. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 6269 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2006. doi: 10.1117/12.669991.

[11] R. D. Wordsworth, F. Forget, F. Selsis, E. Millour, B. Charnay, and J.-B. Madeleine. Gliese 581d is the First Discovered Terrestrial-mass Exoplanet in the Habitable Zone. ApJ, 733:L48, June 2011. doi: 10.1088/2041-8205/733/2/L48.

[12] F. Fressin, G. Torres, J. F. Rowe, D. Charbonneau, L. A. Rogers, S. Ballard, N. M. Batalha, W. J. Borucki, S. T. Bryson, L. A. Buchhave, D. R. Ciardi, J.-M. Désert, C. D. Dressing, D. C. Fabrycky, E. B. Ford, T. N. Gautier, III, C. E. Henze, M. J. Holman, A. Howard, S. B. Howell, J. M. Jenkins, D. G. Koch, D. W. Latham, J. J. Lissauer, G. W. Marcy, S. N. Quinn, D. Ragozzine, D. D. Sasselov, S. Seager, T. Barclay, F. Mullally, S. E. Seader, M. Still, J. D. Twicken, S. E. Thompson, and K. Uddin. Two Earth-sized planets orbiting Kepler-20. Nature, 482:195–198, February 2012. doi: 10.1038/nature10780.

[13] M. J. Holman and N. W. Murray. The Use of Transit Timing to Detect Terrestrial-Mass Extrasolar Planets. *Science*, 307:1288–1291, February 2005. doi: 10.1126/science.1107822.

[14] I. A. G. Snellen. A new method for probing the atmospheres of transiting exoplanets. MNRAS, 353:L1–L6, September 2004. doi: 10.1111/j.1365-2966.2004.08169.x.

[15] D. M. Kipping, G. Á. Bakos, L. Buchhave, D. Nesvorný, and A. Schmitt. The Hunt for Exomoons with Kepler (HEK). I. Description of a New Observational project. ApJ, 750:115, May 2012. doi: 10.1088/0004-637X/750/2/115.

[16] J. Wambsganss. Microlensing Surveys in Search of Extrasolar Planets. In J. Beaulieu, A. Lecavelier Des Etangs, and C. Terquem, editors, *Extrasolar Planets: Today and Tomorrow*, volume 321 of *Astronomical Society of the Pacific Conference Series*, page 47, December 2004.

[17] J.-P. Beaulieu, D. P. Bennett, P. Fouqué, A. Williams, M. Dominik, U. G. Jørgensen, D. Kubas, A. Cassan, C. Coutures, J. Greenhill, K. Hill, J. Menzies, P. D. Sackett, M. Albrow, S. Brillant, J. A. R. Caldwell, J. J. Calitz, K. H. Cook, E. Corrales, M. Desort, S. Dieters, D. Dominis, J. Donatowicz, M. Hoffman, S. Kane, J.-B. Marquette, R. Martin, P. Meintjes, K. Pollard, K. Sahu, C. Vinter, J. Wambsganss, K. Woller, K. Horne, I. Steele, D. M. Bramich, M. Burgdorf, C. Snodgrass, M. Bode, A. Udalski, M. K. Szymański, M. Kubiak, T. Więckowski, G. Pietrzyński, I. Soszyński, O. Szewczyk, Ł. Wyrzykowski, B. Paczyński, F. Abe, I. A. Bond, T. R. Britton, A. C. Gilmore, J. B. Hearnshaw, Y. Itow, K. Kamiya, P. M. Kilmartin, A. V.

Korpela, K. Masuda, Y. Matsubara, M. Motomura, Y. Muraki, S. Nakamura, C. Okada, K. Ohnishi, N. J. Rattenbury, T. Sako, S. Sato, M. Sasaki, T. Sekiguchi, D. J. Sullivan, P. J. Tristram, P. C. M. Yock, and T. Yoshioka. Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. Nature, 439:437–440, January 2006. doi: 10.1038/nature04441.

[18] G. Ingrosso, S. C. Novati, F. de Paolis, P. Jetzer, A. A. Nucita, and A. F. Zakharov. Pixel lensing as a way to detect extrasolar planets in M31. MNRAS, 399:219–228, October 2009. doi: 10.1111/j.1365-2966.2009.15184.x.

[19] G. Marcy, R. P. Butler, D. Fischer, S. Vogt, J. T. Wright, C. G. Tinney, and H. R. A. Jones. Observed Properties of Exoplanets: Masses, Orbits, and Metallicities. *Progress of Theoretical Physics Supplement*, 158:24–42, 2005. doi: 10.1143/PTPS.158.24.

[20] R. P. Butler, J. T. Wright, G. W. Marcy, D. A. Fischer, S. S. Vogt, C. G. Tinney, H. R. A. Jones, B. D. Carter, J. A. Johnson, C. McCarthy, and A. J. Penny. Catalog of Nearby Exoplanets. ApJ, 646:505–522, July 2006. doi: 10.1086/504701.

[21] F. Fressin, G. Torres, D. Charbonneau, S. T. Bryson, J. Christiansen, C. D. Dressing, J. M. Jenkins, L. M. Walkowicz, and N. M. Batalha. The false positive rate of Kepler and the occurrence of planets. *ArXiv e-prints*, January 2013.

[22] N. M. Batalha, J. F. Rowe, S. T. Bryson, T. Barclay, C. J. Burke, D. A. Caldwell, J. L. Christiansen, F. Mullally, S. E. Thompson, T. M. Brown, A. K. Dupree, D. C. Fabrycky, E. B. Ford, J. J. Fortney, R. L. Gilliland, H. Isaacson, D. W. Latham, G. W. Marcy, S. N. Quinn, D. Ragozzine, A. Shporer, W. J. Borucki, D. R. Ciardi, T. N. Gautier, III, M. R. Haas, J. M. Jenkins, D. G. Koch, J. J. Lissauer, W. Rapin, G. S. Basri, A. P. Boss, L. A. Buchhave, J. A. Carter, D. Charbonneau, J. Christensen-Dalsgaard, B. D. Clarke, W. D. Cochran, B.-O. Demory, J.-M. Desert, E. Devore, L. R. Doyle, G. A. Esquerdo, M. Everett, F. Fressin, J. C. Geary, F. R. Girouard, A. Gould, J. R. Hall, M. J. Holman, A. W. Howard, S. B. Howell, K. A. Ibrahim, K. Kinemuchi, H. Kjeldsen, T. C. Klaus, J. Li, P. W. Lucas, S. Meibom, R. L. Morris, A. Prša, E. Quintana, D. T. Sanderfer, D. Sasselov, S. E. Seader, J. C. Smith, J. H. Steffen, M. Still, M. C. Stumpe, J. C. Tarter, P. Tenenbaum, G. Torres, J. D. Twicken, K. Uddin, J. Van Cleve, L. Walkowicz, and W. F. Welsh. Planetary Candidates Observed by Kepler. III. Analysis of the First 16 Months of Data. ApJS, 204:24, February 2013. doi: 10.1088/0067-0049/204/2/24.

[23] E. Chiang and G. Laughlin. The Minimum-Mass Extrasolar Nebula: In-Situ Formation of Close-In Super-Earths. *ArXiv e-prints*, November 2012.

[24] D. C. Fabrycky, J. J. Lissauer, D. Ragozzine, J. F. Rowe, E. Agol, T. Barclay, N. Batalha, W. Borucki, D. R. Ciardi, E. B. Ford, J. C. Geary, M. J. Holman,

J. M. Jenkins, J. Li, R. C. Morehead, A. Shporer, J. C. Smith, J. H. Steffen, and M. Still. Architecture of Kepler's Multi-transiting Systems: II. New investigations with twice as many candidates. *ArXiv e-prints*, February 2012.

[25] M. Jurić and S. Tremaine. Dynamical Origin of Extrasolar Planet Eccentricity Distribution. ApJ, 686:603–620, October 2008. doi: 10.1086/590047.

[26] D. N. C. Lin and S. Ida. On the Origin of Massive Eccentric Planets. ApJ, 477:781, March 1997. doi: 10.1086/303738.

[27] D. A. Fischer and J. Valenti. The Planet-Metallicity Correlation. ApJ, 622: 1102–1117, April 2005. doi: 10.1086/428383.

[28] J. I. González Hernández, G. Israelian, N. C. Santos, S. Sousa, E. Delgado-Mena, V. Neves, and S. Udry. Chemical Abundances of Volatiles and Refractories in Solar Analogs with and without Planets: No Relation with Terrestial Planets. In C. Johns-Krull, M. K. Browning, and A. A. West, editors, *16th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun*, volume 448 of *Astronomical Society of the Pacific Conference Series*, page 879, December 2011.

[29] L. Kaltenegger, A. Segura, and S. Mohanty. Model Spectra of the First Potentially Habitable Super-Earth Gl581d. ApJ, 733:35, May 2011. doi: 10.1088/0004-637X/733/1/35.

[30] K. Heng. The Study of Climate on Alien Worlds. *ArXiv e-prints*, June 2012.

[31] X. Dumusque, F. Pepe, C. Lovis, D. Ségransan, J. Sahlmann, W. Benz, F. Bouchy, M. Mayor, D. Queloz, N. Santos, and S. Udry. An Earth-mass planet orbiting α Centauri B. Nature, 491:207–211, November 2012. doi: 10.1038/nature11572.

[32] F. Pepe, C. Lovis, D. Ségransan, W. Benz, F. Bouchy, X. Dumusque, M. Mayor, D. Queloz, N. C. Santos, and S. Udry. The HARPS search for Earth-like planets in the habitable zone. I. Very low-mass planets around ¡ASTROBJ¿HD 20794¡/ASTROBJ¿, ¡ASTROBJ¿HD 85512¡/ASTROBJ¿, and ¡ASTROBJ¿HD 192310¡/ASTROBJ¿. A&A, 534:A58, October 2011. doi: 10.1051/0004-6361/201117055.

[33] M. Tuomi, G. Anglada-Escudé, E. Gerlach, H. R. A. Jones, A. Reiners, E. J. Rivera, S. S. Vogt, and R. P. Butler. Habitable-zone super-Earth candidate in a six-planet system around the K2.5V star HD 40307. A&A, 549:A48, January 2013. doi: 10.1051/0004-6361/201220268.

[34] J. F. Kasting, D. P. Whitmire, and R. T. Reynolds. Habitable Zones around Main Sequence Stars. Icarus, 101:108–128, January 1993. doi: 10.1006/icar.1993.1010.

[35] F. C. Adams and G. Laughlin. Long-Term Evolution of Close Planets Including the Effects of Secular Interactions. ApJ, 649:1004–1009, October 2006. doi: 10.1086/506145.

[36] J. Laskar, F. Joutel, and P. Robutel. Stabilization of the earth's obliquity by the moon. Nature, 361:615–617, February 1993. doi: 10.1038/361615a0.

[37] R. Lathe. Fast tidal cycling and the origin of life. Icarus, 168:18–22, March 2004. doi: 10.1016/j.icarus.2003.10.018.

[38] P. Ward and D. Brownlee. *Rare earth : why complex life is uncommon in the universe*. 2000.

[39] I. Kant. *Allgemeine Naturgeschichte und Theorie des Himmels*. 1755.

[40] P.-S. Marquis de Laplace. *Exposition du systeme du monde*. 1798.

[41] J. H. Jeans. The part played by rotation in cosmic evolution. MNRAS, 77: 186–199, January 1917.

[42] W. H. McCrea. The Origin of the Solar System. *Royal Society of London Proceedings Series A*, 256:245–266, June 1960. doi: 10.1098/rspa.1960.0108.

[43] V. S. Safronov. *Evoliutsiia doplanetnogo oblaka.* 1969.

[44] M. J. McCaughrean and C. R. O'dell. Direct Imaging of Circumstellar Disks in the Orion Nebula. AJ, 111:1977, May 1996. doi: 10.1086/117934.

[45] J. I. Lunine, D. P. O'Brien, S. N. Raymond, A. Morbidelli, T. Quinn, and A. Graps. Dynamical Models of Terrestrial Planet Formation. *ArXiv e-prints*, June 2009.

[46] A. Morbidelli, J. I. Lunine, D. P. O'Brien, S. N. Raymond, and K. J. Walsh. Building Terrestrial Planets. *Annual Review of Earth and Planetary Sciences*, 40:251–275, May 2012. doi: 10.1146/annurev-earth-042711-105319.

[47] J. J. Lissauer. Planet formation. ARA&A, 31:129–174, 1993. doi: 10.1146/annurev.aa.31.090193.001021.

[48] P. J. Armitage. *Astrophysics of Planet Formation*. 2010.

[49] J. S. Lewis. *Physics and chemistry of the solar system*. 1997.

[50] C. Hayashi. Structure of the Solar Nebula, Growth and Decay of Magnetic Fields and Effects of Magnetic and Turbulent Viscosities on the Nebula. *Progress of Theoretical Physics Supplement*, 70:35–53, 1981. doi: 10.1143/PTPS.70.35.

[51] L. Hartmann. *Accretion Processes in Star Formation*. June 1998.

[52] S. J. Weidenschilling. Dust to planetesimals - Settling and coagulation in the solar nebula. Icarus, 44:172–189, October 1980. doi: 10.1016/0019-1035(80)90064-0.

[53] C. P. Dullemond and C. Dominik. Dust coagulation in protoplanetary disks: A rapid depletion of small grains. A&A, 434:971–986, May 2005. doi: 10.1051/0004-6361:20042080.

[54] S. J. Weidenschilling. Aerodynamics of solid bodies in the solar nebula. MNRAS, 180:57–70, July 1977.

[55] T. Birnstiel, C. P. Dullemond, and F. Brauer. Gas- and dust evolution in protoplanetary disks. A&A, 513:A79, April 2010. doi: 10.1051/0004-6361/200913731.

[56] F. Windmark, T. Birnstiel, C. W. Ormel, and C. P. Dullemond. Breaking through: The effects of a velocity distribution on barriers to dust growth. A&A, 544:L16, August 2012. doi: 10.1051/0004-6361/201220004.

[57] N. Haghighipour and A. P. Boss. On Gas Drag-Induced Rapid Migration of Solids in a Nonuniform Solar Nebula. ApJ, 598:1301–1311, December 2003. doi: 10.1086/378950.

[58] W. K. M. Rice, G. Lodato, J. E. Pringle, P. J. Armitage, and I. A. Bonnell. Accelerated planetesimal growth in self-gravitating protoplanetary discs. MNRAS, 355:543–552, December 2004. doi: 10.1111/j.1365-2966.2004.08339.x.

[59] H. H. Klahr and T. Henning. Particle-Trapping Eddies in Protoplanetary Accretion Disks. Icarus, 128:213–229, July 1997. doi: 10.1006/icar.1997.5720.

[60] P. Goldreich and W. R. Ward. The Formation of Planetesimals. ApJ, 183:1051–1062, August 1973. doi: 10.1086/152291.

[61] A. Toomre. On the gravitational stability of a disk of stars. ApJ, 139:1217–1238, May 1964. doi: 10.1086/147861.

[62] J. N. Cuzzi, A. R. Dobrovolskis, and J. M. Champney. Particle-gas dynamics in the midplane of a protoplanetary nebula. Icarus, 106:102, November 1993. doi: 10.1006/icar.1993.1161.

[63] S. J. Weidenschilling. Can gravitation instability form planetesimals? Icarus, 116:433–435, August 1995. doi: 10.1006/icar.1995.1136.

[64] A. N. Youdin and J. Goodman. Streaming Instabilities in Protoplanetary Disks. ApJ, 620:459–469, February 2005. doi: 10.1086/426895.

[65] A. Johansen, J. S. Oishi, M.-M. Mac Low, H. Klahr, T. Henning, and A. Youdin. Rapid planetesimal formation in turbulent circumstellar disks. Nature, 448:1022–1025, August 2007. doi: 10.1038/nature06086.

[66] G. R. Stewart and G. W. Wetherill. Evolution of planetesimal velocities. Icarus, 74:542–553, June 1988. doi: 10.1016/0019-1035(88)90120-0.

[67] S. Ida and J. Makino. N-body simulation of gravitational interaction between planetesimals and a protoplanet. II - Dynamical friction. Icarus, 98:28–37, July 1992. doi: 10.1016/0019-1035(92)90203-J.

[68] R. Greenberg, W. K. Hartmann, C. R. Chapman, and J. F. Wacker. Planetesimals to planets - Numerical simulation of collisional evolution. Icarus, 35: 1–26, July 1978. doi: 10.1016/0019-1035(78)90057-X.

[69] G. W. Wetherill and G. R. Stewart. Accumulation of a swarm of small planetesimals. Icarus, 77:330–357, February 1989. doi: 10.1016/0019-1035(89) 90093-6.

[70] S. Ida and J. Makino. Scattering of planetesimals by a protoplanet - Slowing down of runaway growth. Icarus, 106:210, November 1993. doi: 10.1006/icar. 1993.1167.

[71] E. Kokubo and S. Ida. Oligarchic Growth of Protoplanets. Icarus, 131:171–178, January 1998. doi: 10.1006/icar.1997.5840.

[72] J. J. Lissauer. Timescales for planetary accretion and the structure of the protoplanetary disk. Icarus, 69:249–265, February 1987. doi: 10.1016/ 0019-1035(87)90104-7.

[73] J. B. Pollack, O. Hubickyj, P. Bodenheimer, J. J. Lissauer, M. Podolak, and Y. Greenzweig. Formation of the Giant Planets by Concurrent Accretion of Solids and Gas. Icarus, 124:62–85, November 1996. doi: 10.1006/icar.1996. 0190.

[74] S. J. Weidenschilling, D. Spaute, D. R. Davis, F. Marzari, and K. Ohtsuki. Accretional Evolution of a Planetesimal Swarm. Icarus, 128:429–455, August 1997. doi: 10.1006/icar.1997.5747.

[75] S. Ida, T. Guillot, and A. Morbidelli. Accretion and Destruction of Planetesimals in Turbulent Disks. ApJ, 686:1292–1301, October 2008. doi: 10.1086/591903.

[76] S. T. Stewart and Z. M. Leinhardt. Velocity-Dependent Catastrophic Disruption Criteria for Planetesimals. ApJ, 691:L133–L137, February 2009. doi: 10.1088/0004-637X/691/2/L133.

[77] J. E. Chambers and G. W. Wetherill. Making the Terrestrial Planets: N-Body Integrations of Planetary Embryos in Three Dimensions. Icarus, 136:304–327, December 1998. doi: 10.1006/icar.1998.6007.

[78] R. Morishima, J. Stadel, and B. Moore. From planetesimals to terrestrial planets: N-body simulations including the effects of nebular gas and giant planets. Icarus, 207:517–535, June 2010. doi: 10.1016/j.icarus.2009.11.038.

[79] L. Dones and S. Tremaine. On the origin of planetary spins. Icarus, 103:67–92, May 1993. doi: 10.1006/icar.1993.1059.

[80] J. J. Lissauer, L. Dones, and K. Ohtsuki. *Origin and Evolution of Terrestrial Planet Rotation*, pages 101–112. 2000.

[81] W. Benz, A. G. W. Cameron, and W. L. Slattery. Planetary Collision Calculations: Origin of the Moon. In *Lunar and Planetary Institute Science Conference Abstracts*, volume 18 of *Lunar and Planetary Inst. Technical Report*, page 60, March 1987.

[82] R. M. Canup. Simulations of a late lunar-forming impact. Icarus, 168:433–456, April 2004. doi: 10.1016/j.icarus.2003.09.028.

[83] R. M. Canup. Lunar-forming collisions with pre-impact rotation. Icarus, 196: 518–538, August 2008. doi: 10.1016/j.icarus.2008.03.011.

[84] A. M. Davis. *Volatile Evolution and Loss*, pages 295–307. 2006.

[85] W. Benz, W. L. Slattery, and A. G. W. Cameron. Collisional stripping of Mercury's mantle. Icarus, 74:516–528, June 1988. doi: 10.1016/0019-1035(88)90118-2.

[86] G. W. Wetherill and G. R. Stewart. Formation of planetary embryos - Effects of fragmentation, low relative velocity, and independent variation of eccentricity and inclination. Icarus, 106:190, November 1993. doi: 10.1006/icar.1993.1166.

[87] H. Levison, D. Nesvorny, C. Agnor, and A. Morbidelli. The Role of Dynamical Friction in Terrestrial Planet Formation. In *AAS/Division for Planetary Sciences Meeting Abstracts #37*, volume 37 of *Bulletin of the American Astronomical Society*, page 666, August 2005.

[88] P. Bodenheimer and J. B. Pollack. Calculations of the accretion and evolution of giant planets The effects of solid cores. Icarus, 67:391–408, September 1986. doi: 10.1016/0019-1035(86)90122-3.

[89] G. P. Kuiper. On the Origin of the Solar System. *Proceedings of the National Academy of Science*, 37:1–14, January 1951. doi: 10.1073/pnas.37.1.1.

[90] A. G. W. Cameron. Physics of the primitive solar accretion disk. *Moon and Planets*, 18:5–40, February 1978. doi: 10.1007/BF00896696.

[91] A. P. Boss. Giant planet formation by gravitational instability. *Science*, 276: 1836–1839, 1997. doi: 10.1126/science.276.5320.1836.

[92] C. F. Gammie. Nonlinear Outcome of Gravitational Instability in Cooling, Gaseous Disks. ApJ, 553:174–183, May 2001. doi: 10.1086/320631.

[93] A. C. Boley. The Two Modes of Gas Giant Planet Formation. ApJ, 695: L53–L57, April 2009. doi: 10.1088/0004-637X/695/1/L53.

[94] A. C. Boley, T. Hayfield, L. Mayer, and R. H. Durisen. Clumps in the outer disk by disk instability: Why they are initially gas giants and the legacy of disruption. Icarus, 207:509–516, June 2010. doi: 10.1016/j.icarus.2010.01.015.

[95] A. P. Boss. Flux-limited Diffusion Approximation Models of Giant Planet Formation by Disk Instability. ApJ, 677:607–615, April 2008. doi: 10.1086/533496.

[96] H. F. Levison and C. Agnor. The Role of Giant Planets in Terrestrial Planet Formation. AJ, 125:2692–2713, May 2003. doi: 10.1086/374625.

[97] S. N. Raymond, T. Quinn, and J. I. Lunine. High-resolution simulations of the final assembly of Earth-like planets I. Terrestrial accretion and dynamics. Icarus, 183:265–282, August 2006. doi: 10.1016/j.icarus.2006.03.011.

[98] H. Levison, M. Duncan, C. Capobianco, and D. Minton. Planetesimal-Driven Migration Verses Type I: Rethinking Giant Planet Formation. In *EPSC-DPS Joint Meeting 2011*, page 1476, October 2011.

[99] K. J. Walsh, A. Morbidelli, S. N. Raymond, D. P. O'Brien, and A. M. Mandell. Populating the asteroid belt from two parent source regions due to the migration of giant planets "The Grand Tack". *Meteoritics and Planetary Science*, 47:1941–1947, December 2012. doi: 10.1111/j.1945-5100.2012.01418.x.

[100] K. Tsiganis, R. Gomes, A. Morbidelli, and H. F. Levison. Origin of the orbital architecture of the giant planets of the Solar System. Nature, 435:459–461, May 2005. doi: 10.1038/nature03539.

[101] S. Grimm and J. Stadel. GENGA: a GPU parallel symplectic integrator for planetesimal dynamics and collisions. *in prep.*, 2013.

[102] J. E. Chambers. A hybrid symplectic integrator that permits close encounters between massive bodies. MNRAS, 304:793–799, April 1999. doi: 10.1046/j.1365-8711.1999.02379.x.

[103] J. Wisdom and M. Holman. Symplectic maps for the n-body problem. AJ, 102:1528–1538, October 1991. doi: 10.1086/115978.

[104] J. G. Stadel. *Cosmological N-body simulations and their analysis*. PhD thesis, UNIVERSITY OF WASHINGTON, 2001.

[105] G. W. Wetherill. Formation of the terrestrial planets. ARA&A, 18:77–113, 1980. doi: 10.1146/annurev.aa.18.090180.000453.

[106] G. W. Wetherill. Formation of the earth. *Annual Review of Earth and Planetary Sciences*, 18:205–256, 1990. doi: 10.1146/annurev.ea.18.050190.001225.

[107] C. B. Agnor, R. M. Canup, and H. F. Levison. On the Character and Consequences of Large Impacts in the Late Stage of Terrestrial Planet Formation. Icarus, 142:219–237, November 1999. doi: 10.1006/icar.1999.6201.

[108] S. N. Raymond, T. Quinn, and J. I. Lunine. Making other earths: dynamical simulations of terrestrial planet formation and water delivery. Icarus, 168: 1–17, March 2004. doi: 10.1016/j.icarus.2003.11.019.

[109] E. Kokubo, J. Kominami, and S. Ida. Formation of Terrestrial Planets from Protoplanets. I. Statistics of Basic Dynamical Properties. ApJ, 642:1131–1139, May 2006. doi: 10.1086/501448.

[110] S. N. Raymond, D. P. O'Brien, A. Morbidelli, and N. A. Kaib. Building the terrestrial planets: Constrained accretion in the inner Solar System. Icarus, 203:644–662, October 2009. doi: 10.1016/j.icarus.2009.05.016.

[111] S. N. Raymond, R. Barnes, and N. A. Kaib. Predicting Planets in Known Extrasolar Planetary Systems. III. Forming Terrestrial Planets. ApJ, 644: 1223–1231, June 2006. doi: 10.1086/503594.

[112] D. P. O'Brien, A. Morbidelli, and H. F. Levison. Terrestrial planet formation with strong dynamical friction. Icarus, 184:39–58, September 2006. doi: 10.1016/j.icarus.2006.04.005.

[113] M. Nagasawa and S. Ida. Sweeping Secular Resonances in the Kuiper Belt Caused by Depletion of the Solar Nebula. AJ, 120:3311–3322, December 2000. doi: 10.1086/316856.

[114] E. Kokubo and H. Genda. Formation of Terrestrial Planets from Protoplanets Under a Realistic Accretion Condition. ApJ, 714:L21–L25, May 2010. doi: 10.1088/2041-8205/714/1/L21.

[115] S. Elser, B. Moore, J. Stadel, and R. Morishima. How common are Earth-Moon planetary systems? Icarus, 214:357–365, August 2011. doi: 10.1016/j.icarus.2011.05.025.

[116] J. C. Bond, D. S. Lauretta, and D. P. O'Brien. Making the Earth: Combining dynamics and chemistry in the Solar System. Icarus, 205:321–337, February 2010. doi: 10.1016/j.icarus.2009.07.037.

[117] S. Elser, M. R. Meyer, and B. Moore. On the origin of elemental abundances in the terrestrial planets. Icarus, 221:859–874, November 2012. doi: 10.1016/j.icarus.2012.09.016.

[118] S. Elser, S. L. Grimm, and J. G. Stadel. Super-Earths and dynamical stability of planetary systems: first parallel GPU simulations using GENGA. MNRAS, 433:2194–2205, August 2013. doi: 10.1093/mnras/stt883.

[119] A. P. Jackson and M. C. Wyatt. Debris from terrestrial planet formation: the Moon-forming collision. MNRAS, 425:657–679, September 2012. doi: 10.1111/j.1365-2966.2012.21546.x.

[120] D. C. Richardson, T. Quinn, J. Stadel, and G. Lake. Direct Large-Scale N-Body Simulations of Planetesimal Dynamics. Icarus, 143:45–59, January 2000. doi: 10.1006/icar.1999.6243.

[121] U. Wiechert, A. N. Halliday, D.-C. Lee, G. A. Snyder, L. A. Taylor, and D. Rumble. Oxygen Isotopes and the Moon-Forming Giant Impact. *Science*, 294:345–348, October 2001. doi: 10.1126/science.1063037.

[122] A. Reufer, M. M. M. Meier, W. Benz, and R. Wieler. A hit-and-run giant impact scenario. Icarus, 221:296–299, September 2012. doi: 10.1016/j.icarus.2012.07.021.

[123] E. Miller-Ricci, M. R. Meyer, S. Seager, and L. Elkins-Tanton. On the Emergent Spectra of Hot Protoplanet Collision Afterglows. ApJ, 704:770–780, October 2009. doi: 10.1088/0004-637X/704/1/770.